

Antirealist Essentialism

Jonathan Edward Banks

The University of Leeds

School of Philosophy, Religion and History of Science

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

August 2014

The candidate confirms that this work is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2014 The University of Leeds and Jonathan Edward Banks.

The right of Jonathan Edward Banks to be identified as the author of this work has been asserted by him in accordance with the Copyright, Designs and Patents act 1988.

Acknowledgements

I would like to thank my sage supervision team, John Divers and Scott Shalkowski, who did an exemplary job moulding a less than exceptional mind into a (hopefully) competent young philosopher. Many hours of discussion have played an incalculable role not only in the production of this thesis, but in my development as a scholar. Thank you to my examiners, Helen Steward and Fabrice Correia, for a stimulating viva and great feedback.

I would also like to thank my fellow grad students both past and present for much help, friendship and lively debate (both philosophical and otherwise). Of particular note are Jordan Bartol, Dominic Berry, Richard Caves, Fran Fairbairn, Nanna Lüders-Kaalund, Kerry McKenzie, Bert Muckle, Dave Race, Efram Sera-Shriar, and Carl ‘Swiggles’ Warom. Many of you helped me out in my many hours of need, with a pint, a place to stay, or just some good company. I assure you, it never went unappreciated.

Extra special thanks go to Dani Adams and Sarah Adams, with whom I have had the distinct pleasure of sharing my philosophical education. These fantastic, clever people have read more of my work, and listened to more of my ramblings, than anyone. Also of extra special note are the contributions of Thomas Brouwer and Michael Bench-Capon, from whom I have probably learnt as much as I have from my supervisors. Much of this thesis owes a great debt to discussions with both of them, from the insights they shared with me, and the amount they taught me (often quite explicitly). You sirs, are both scholars and gentlemen.

I also received much help and support from the Philosophy department at Leeds. The department awarded me a fees scholarship; Nick Jones offered me work as his assistant; and the office staff were ever-helpful. I’d like to thank the faculty at Leeds for creating such an engaging research environment. In particular I’m grateful to Elizabeth Barnes, Ross Cameron, Nick Jones, Robin Le Poidevin, Heather Logue, Andrew McGonigal, Aaron Meskin, Jason Turner, and Robbie

Williams for many useful discussions of my research, life as an academic, and how to be a better teacher. I'd also like to thank the late, great Nick Maloney of St Boniface's College, Plymouth. He introduced me to philosophy with humour and candour. I never got a chance to properly thank him in person.

I'd like to thank all of the students I've taught over the past five years, even the rubbish ones. Without the challenges and inspiration I encountered whilst teaching I would not have developed the passion for education that I hope will keep me striving to become a better teacher.

To all of my friends outside of the academic bubble: I'm sorry for being a boring recluse for the last five years, and I'm sorry for how much you're going to have to put up with me from now on.

My wonderful family gave me a huge amount of support, both emotional and monetary. Thank you for always having faith in my abilities, and unwaveringly supporting my passion.

My final and greatest thanks go to Michaela Livingstone. She is my favourite person. She kept me sane when I would otherwise have struggled. She makes a great cup of tea. She is way better than Philosophy.

Abstract

This project is an investigation into the prospects for an antirealist theory of essence. Essentialism is the claim that at least some things have some of their properties essentially. Essentialist discourse includes claims such as “Socrates is essentially human”, and “Socrates is accidentally bearded”. Historically, there are two ways of interpreting essentialist discourse. I call these positions ‘modal essentialism’ and ‘neo-Aristotelian essentialism’. According to modal essentialism, for Socrates to be essentially human is for it to be necessary that he be a human if he exists, and for Socrates to accidentally have a beard is for it to be contingent that Socrates has a beard if he exists. According to neo-Aristotelian essentialism, objects have definitions in something like the way words do. For Socrates to be essentially human but accidentally bearded is for it to be part of the definition of Socrates that he is human, but not part of that definition that he is bearded. I argue that both are susceptible to antirealist interpretation. This thesis sets about showing that this is the case.

In Chapters One and Two I investigate neo-conventionalist theories of modality, in the hope of using such a position to develop an antirealist modal essentialism. In Chapter Three I discuss the debate between modal and neo-Aristotelian essentialism and conclude that it is by no means settled. In Chapter Four I develop an antirealist neo-Aristotelian essentialism based on the mechanism of one of the neo-conventionalist accounts of modality. In Chapter Five I argue that this account is in a better position to give an essentialist theory of necessity than its realist counterparts. I conclude that, regardless of whether one is a modal or neo-Aristotelian essentialist, antirealist essentialism is a viable theory of essence that is worthy of consideration in contemporary debate.

Contents

Acknowledgements.....	vi
Abstract	viii
Introduction	1
1. Introduction.....	1
2. Project Goals.....	4
3. Essentialism	5
4. Antirealism	9
5. Conventionalism	10
5.1 Traditional Conventionalism.....	10
5.2 The Downfall of Traditional Conventionalism	13
5.3 What Would a New Conventionalist Account Need to be Like?	16
5.4 Conventionalism vs. Antirealism.....	19
6. Chapter Outline.....	21
Chapter One: Deflationary and Humean Neo-Conventionalism.....	26
1. Introduction.....	26
2. Neo-Conventionalism	27
2.1 Deflationary and Humean Neo-Conventionalism.....	27
2.2 A Note on Naturalness	30
3. Truth by Convention	33
3.1 Does Neo-Conventionalism Rely on Truth by Convention?	33
3.2 Does Type 2 Grounding Count as Truth by Convention?	37
3.3 Is Reliance on Truth by Convention Problematic for the Deflationist?	39
4. The Contingency Problem	42

4.1 Setting up the Problem	42
4.2 The Einheuser Framework	47
4.3. Applying the Einheuser Framework to Neo-Conventionalism	51
4.4 The Contingency Problem: Continued	53
4.5 An Unsuccessful Response	56
4.6 The Proposed Solution	57
4.7 Guarding Against the Original Contingency Threat	59
5. Two Unanswered Questions for Deflationism and Humeanism	61
6. Conclusion	63
Chapter Two: Sidelleian Conventionalism	64
1. Introduction	64
2. Sidelleian Conventionalism	64
3. General Principles of Individuation	67
4. The Link between Modality and Ontology	72
5. How Does this Give us Essence?	75
6. Advantages Over Other Neo-Conventionalisms	76
7. The Indexicality Problem	77
8. Responding to the Indexicality Problem	82
9. Reliance on Analyticity	84
9.1. The Analyticity of the GPIs	85
9.2. Rejecting the Analyticity of the GPIs	86
9.3 Renouncing Analyticity Altogether	88
10. Conclusion	93
Chapter Three: Fine on Essence	94

1. Introduction.....	94
1.2 Fine's Characterisation of Essence	96
2. Against Modal Essentialism.....	98
3. Theory of Essence.....	102
3.1 Characterisation of Essence as Definition	102
3.2 True in Virtue of the Nature of... ..	105
3.3 Predicational and Sentential Forms.....	106
3.4 The Extent of Essence.....	107
3.4.1 Consequential/Constitutive Essence	107
3.4.2 The Role of Logical Concepts.....	110
3.4.3 Reviving Constitutive Essence	112
3.4.4 Mediate/Immediate Essence.....	113
3.5 Possible Sources of Regress.....	114
3.5.1 Reflexive Essence	114
3.5.2 Reciprocal Essence	115
3.6 Essential Manifold	116
4. Responding to Fine.....	118
4.1 Case 3	118
4.2 Case 2	123
4.3 Case 4	125
4.4 Case 1	129
5. Conclusion.....	130
Chapter Four: Antirealism about Neo-Aristotelian Essence.....	132
1. Introduction.....	132

2. Neo-Aristotelian Essentialism.....	133
3. Options for Antirealism.....	134
3.1 No-Reduction.....	136
3.2 Robust-TIVON No-Objects.....	137
3.3 No-TIVON Robust-Objects.....	138
3.4 No-TIVON No-Objects.....	139
4. Conventional Mechanism.....	140
4.1 No-TIVON No-Objects.....	141
4.2 No-TIVON Robust-Objects.....	144
4.3 Investigating the Type 2 No-TIVON Robust-Objects Option	146
4.4 Investigating the Type 3 No-TIVON Robust-Objects Option	147
4.5 Can (Robust) Objects Exist Without (Robust) Essential Profiles?	148
4.6 Can (Robust) Kinds Exist Without (Robust) Essential Profiles?	149
4.7 If the answers to one or both of these questions is ‘yes’ then what is the theoretical role of essence?.....	152
4.8 If the answers to one or both of the first two questions is ‘yes’, then why do we essentialise?	152
5. The Proposed Account.....	153
5.1 Kinds.....	156
5.2 Robust-Objects or No-Objects?.....	160
5.3 Dispensing with the Propositional Formulation of Essentialist Predication.....	160
6. Generic Essence	163
6.1 Realism and Generic Essence.....	168

6.2 Antirealism and Generic Essence	170
7. Conclusion.....	172
Chapter Five: Essentialist Theories of Necessity.....	174
1. Introduction.....	174
2. Fine	175
2.1 Finean Essentialism	175
2.2 The Finean Essentialist Theory of Modality	176
2.3 Finean Logic of Essence	177
2.4 Correia on Essence and Modality	178
2.5 Consequentialist Account.....	180
2.6 Rule-Based Account	182
2.7 Logical and Conceptual Necessity	184
2.8 The Use of Concepts	186
3. Hale's Essentialist Theory of Necessity	187
3.1 Hale on Essence.....	188
3.2 Hale on Necessity.....	191
4. Complications with the Theory of Necessity	192
4.1 Sufficient Grounds	192
4.1.2 A First Response	194
4.1.3 A Second Response	195
4.2 Further complications with the Theories of Necessity	196
4.2.1 Hale to the Rescue?.....	199
4.2.2 Back to Pegasus.....	205
4.3 Reductive Ambitions	206

5. The Choice of Essentialist Theories of Necessity.....	208
6. An Antirealist Essentialist Theory of Necessity	210
6.1 Theory of Necessity	210
6.2 Concepts	210
6.3 Derivative Essence	212
6.4 Commitment to Mere Possibilia	213
6.5 Reductive Ambitions	216
7. Conclusion.....	217
Conclusion.....	218
1. Goals.....	218
2. Future work	220
3. Conclusion.....	222
Bibliography.....	224

Antirealist Essentialism:

An Exercise in Eating Your Cake Without Having it

Introduction

1. Introduction

Essence abounds! The language of our everyday lives and academic investigations is replete with expressions of essentialist beliefs. These beliefs are part ordinary, and part metaphysical in the language we use to express them. They include claims about individuals, like (more ordinarily) “Socrates is accidentally bearded”, and (more metaphysically) “Socrates is essentially human”, and claims about kinds, like “water is accidentally the substance that fills our lakes and rivers”, and “water essentially has the elemental composition H_2O ”.

There are two further distinctions that are adjacent to this one. These too are to be found in the language we use to express everyday claims, and the claims that we use to express each might be (and, indeed, sometimes are) used casually or pre-theoretically as practical synonyms for the claims above. First, we often distinguish between *what* something is and *how* it is. For example, we can distinguish between what Socrates is (a human, a mammal, a material object), and how he is (bearded, a philosopher, the teacher of Plato). Second, we distinguish between those qualities a particular object might or might not have had, and those qualities it must have. For example, Socrates might have been a farmer, or clean shaven; he might never have met Plato, but he could not have been a trumpet, nor could he have been born to different parents.

This project is concerned with first, how we should understand these different kinds of claims and how they relate to each other (that is, which, if any, of the casual claims of synonymy between expressions of these related distinctions hold (and what

philosophy underpins that casual synonymy)), and second, what successfully making such claims commits us to.

Historically, philosophers have favoured one of two ways of understanding the distinction between the essential and the accidental. On the one hand, the historically more popular view is to understand the distinction modally (that is, in terms of the second adjacent distinction). For Socrates to be essentially human is for it to be necessary that he be a human if he exists, and for Socrates to accidentally have a beard is for it to be contingent that Socrates has a beard if he exists.¹ On the other hand, the increasingly popular view of the late 20th and early 21st century is that the distinction should be understood in terms of *what it is to be* a specific thing, or of a specific kind (this is in line with the first adjacent distinction). If we think of things as having definitions in something like the way words do, then for Socrates to be essentially human but accidentally bearded is for it to be part of the definition of Socrates that he is human, but not part of that definition that he is bearded. Hereafter I refer to the former position as *modal essentialism*, and the latter as *neo-Aristotelian essentialism*.²

The conflict between modal and neo-Aristotelian essentialism features significantly in this project, but it is not the primary focus. My primary focus is on what, once the dust has settled and we have decided what essentialist claims mean, we can conclude about the world around us, given that we do successfully make essentialist claims. When we make true essentialist claims, are we successfully describing the way the world is, independent of us, or are we revealing an aspect of how we represent the world? Call the first option *realism*. Realism about essence claims that, regardless of how we interpret essentialist claims, when we make them

¹ The 'if he exists' may seem redundant in the case of Socrates being accidentally bearded, but it is not. If Socrates is contingent, but has a beard in every world where he exists, then it will still be contingent that Socrates has a beard (because it will not be the case that Socrates has a beard in every possible world) and so he will be accidentally bearded, but also essentially bearded (because he will necessarily have a beard if he exists).

² Neo-Aristotelian because it makes use of *real definitions*, a notion that is often attributed to Aristotle. Cf. Fine (1994), p.2-3, Hale (2013), p.152 fn.17

successfully the source of that success is in our accurately describing reality.³ The alternative position is *antirealism* about essence. Antirealism claims that there are no essentialist facts independent of us. When we make essentialist claims, regardless of what we mean by them, the source of that success is not solely in the world, but (at least in part) in the way that we represent it. The thesis I defend is that, regardless of whether one is a modal essentialist or a neo-Aristotelian essentialist, there is a viable antirealist position available that is at least as favourable as its realist counterparts and so is worthy of discussion in modern debates about essence.

In this introduction I prepare the ground for the work ahead. I establish the limits of the project and suggest how I will go about arguing for the general thesis. In Section Two I state the goals of this project in contrast with questions that, whilst relevant, are beyond the scope of this project to address. In Section Three I say a little more about essentialism, and the rivalry between modal and neo-Aristotelian essentialism. I suggest how one might interpret the debate; an interpretation that I think has a significant effect on what conclusions we can draw from the recent literature. In Section Four I make some comments concerning antirealism about essence. I consider what the antirealist claim (as considered in this project) might amount to, and take some steps to motivate an investigation of antirealism. In Section Five I discuss the primary form of antirealism that I will be discussing in this project, conventionalism. I explain what conventionalism is and why traditional conventionalism failed, and I give conditions that any new conventionalist theory must meet. Whilst conventionalism is a central part of this project, and I will largely be using conventionalist methodology, there are important differences between conventionalism and the antirealism this project supports. I highlight these differences here. In Section Six I give an outline of the project and a breakdown of each chapter.

³ The truth of essentialist claims goes beyond the minimally conceived truth predicate of Wright's (1992) antirealism.

2. Project Goals

This project has three primary goals. The first goal is to show that in the debate between modal essentialism and neo-Aristotelian essentialism, the Finean arguments are not decisive against the modal essentialist, but are unintentionally suggestive of a more convincing reason to favour neo-Aristotelian essentialism. The second goal is to demonstrate that an antirealist theory of essence is a viable position to hold, and one worthy of wider consideration in contemporary metaphysics. The third goal is to show that this is the case regardless of whether one interprets essence as the modal essentialist or the neo-Aristotelian essentialist does.

In the course of this investigation it will become clear that there are some issues that arise that I will have very little to say about. The scope and size of this project are both limited. As such, some peripheral topics that are no less worthy of discussion will go unaddressed. For the sake of maintaining the focus of the project, and not taking meandering detours away from its primary objectives, where possible, I will avoid entering into the following debates:

- The metaphysics of kinds
- The metaphysics of properties
- What things are essentially what
- The exact nature of conventions (whilst I do discuss this, I do not do so in the fulldepth that it warrants)
- The history of essentialist theorising pre-20th century

I will, at times, liberally use terms such as ‘properties’ and ‘facts’ that often come with associated realist metaphysics. I do so for one of two reasons: because I am discussing someone else’s work and I am following their terminology, or to help the discussion flow for the ease of the reader. Unless explicitly stated otherwise or obvious from context, when I use such terms I intend them in a neutral sense that does not endorse any associated metaphysical theory.

There are also some more significant discussions that I shall avoid. Beyond the discussion of it in Section Five of this introduction I will have little to say about truth

by convention, otherwise known as truth in virtue of meaning, or analyticity. Having given a preliminary discussion of the widespread repudiation of truth by convention and the failure of traditional conventionalism about modality, it is an assumption of this thesis that such a notion is defunct and therefore not available as a tool for the antirealist about essence. This is by no means a settled debate.⁴ However, for the sake of this project I will treat it as such, for a full examination would be too significant a detour. This has a significant consequence on how the project will proceed. The first major task of this thesis is to establish an antirealist account of essence that, whilst conventional in nature, does not rely on the much maligned truth by convention.

3. Essentialism

Essentialism owes much of its 20th century revival to Saul Kripke and Hilary Putnam.⁵ The advent of the necessary a posteriori that came as a result of work such as theirs made what was once considered a ‘metaphysical jungle’ into a mainstream subject.⁶ These claims, whilst metaphysically necessarily, are epistemically contingent. Examples include theoretical identities like ‘Water is H₂O’ and ‘heat is molecular motion’, object identities such as ‘Hesperus is Phosphorus’ and essential predications such as ‘Socrates is the son of Phaenarete and Sophroniscus’. The significance of the necessary a posteriori comes primarily from how examples of it motivate essentialism in a way that other examples fail to do. To show this it is worth distinguishing between two kinds of examples that are given in essentialist contexts.

Essentialist claims appear to be concerned with questions of identity. However, there are two closely related but distinct questions here: a modal question of identity, and a temporal one. Within a world *w*, there is the temporal question of diachronic identity, which asks about the same thing over time, and there is the modal question, which asks about the same thing across possible worlds. The temporal question

⁴ For a recent defence of analyticity, see Russell (2008)

⁵ Kripke (1971, 1980), Putnam (1973, 1975). Contemporary modal essentialists include Mackie (2006), and Wildman (2013, forthcoming).

⁶ As called by Quine (1953), p.397

concerns what kinds of change something can survive. Consider Socrates in w_1 at t_1 . We can ask if Socrates still exists in w_1 at t_2 (that is, if Socrates in w_1 at t_1 is identical to Socrates in w_1 at t_2). For example, we think that Socrates can survive a haircut but not being turned into sausages. Conversely, we think that the aggregate of biological matter that makes up Socrates survives being turned into sausages, but not the haircut. The modal question asks if Socrates in w_1 is identical to the Socrates in another world. For example, we think that Socrates in w_1 who is a philosopher is identical to Socrates in w_2 who is a farmer.⁷ However, when invited to consider w_3 where (allegedly) Socrates is a trumpet, we don't think Socrates in w_1 is identical to Socrates in w_3 .

It is important to note that those following Quine in wanting to avoid the 'metaphysical jungle' of essentialism can accept that we can make sense of diachronic identity, even whilst denying that any sense can be made of modal identity. So, whilst the question of diachronic identity may well be closely connected to essentialism, examples of it are not fit to motivate its adoption.⁸ As such, by popularising the necessary a posteriori, the works of Kripke and Putnam provide examples unacceptable to the Quinean sceptic in a way that temporal cases (for example) cannot. The kind of essentialism considered therein is identified with a kind of weak necessity, whereby something is essentially F iff it is F in every world in which it exists. Call this modal essentialism.

Fine challenges modal essentialism, presenting a series of arguments that it does not accurately capture the meaning behind essentialist discourse.⁹ Fine characterises what I call neo-Aristotelian essentialism along the lines of real definition. The

⁷ How one interprets such an identity claim is subject to how one interprets world talk. I will not enter into that discussion here.

⁸ Whilst this may be the case, it does seem that when we consider whether something could survive some temporal change, it is our beliefs about the essence of that thing that we are appealing to when we make our decision. As such, even though we cannot use temporal examples to *motivate* the adoption of essentialism, once we have adopted it we may accept them as an appropriate *application* of it.

⁹ Fine (1994)

essence of an object is the way that thing must be *if it is to be the thing it is*, with the italicised phrase being of central importance to Fine's understanding of essence.¹⁰

Fine's arguments against modal essentialism proceed by giving examples of necessity that do not satisfy Fine's understanding of essence. For example, it is necessary that Socrates is a member of the singleton set {Socrates} if he exists but, according to Fine, Socrates is not essentially a member of {Socrates}, even though it is essential to {Socrates} to have Socrates as a member. This kind of asymmetry is an important part of why Fine thinks modal essentialism is extensionally infelicitous with our understanding of essence. There are several attempts at replying to Fine's arguments, some more successful than others.¹¹ I contest that these responses give Fine's arguments too much credit. Here I suggest an alternative way that we should construe Fine's arguments. We should not see them as a decisive attack on modal essentialism, but as definitive of a more interesting and metaphysically useful way of understanding essence.

I claim that we should interpret the debate between the modal essentialist and the neo-Aristotelian essentialist not as between competing metaphysical explanations of a well understood pre-theoretical notion, but as between two rival notions of essence intended to fit our incomplete pre-philosophical essentialist opinions. Under the common interpretation of the debate, the body of data (comprising our pre-philosophical essentialist opinions) is more extensive and less negotiable. I suggest that the body of data is less extensive and more negotiable, especially when discussing some of the more unusual entities that are the subject of metaphysical debate. Our pre-philosophical essentialist opinions are sufficient to inform us concerning everyday examples (such as whether Socrates could have been born to different parents), but are not sufficient to adequately inform us in more esoteric examples, such as questions concerning the unusual objects of metaphysics (like sets, properties, and numbers). The more obscure and theoretical the subject matter, the less our

¹⁰ Fine develops this position over a series of papers (1994, 1995a, 1995b, 1995c, 2000, 2007). Other notable neo-Aristotelian essentialists include Correia (2006, 2012), Hale (2013), Lowe (2008), Oderberg (2007) Shalkowski (2008).

¹¹ Cf. Correia (2007), Della Rocca (1996), Gorman (2005), Wildman (2013, Forthcoming), Zalta (2006).

opinions inform us, and the more negotiable the results of those opinions are. As such, as soon as we start to introduce philosophically informed intuitions, or even philosophical examples, we run the risk of begging the question by assuming an account of essence to ‘fill in the gaps’ in our essentialist beliefs and applying it to examples from outside our original body of data (our pre-philosophical essentialist opinions). The very fact that there is disagreement in some cases, and (I contend) a profound lack of opinion (from people outside of academia, at least) in other cases should suggest that our understanding of essence is less extensive and more negotiable than the debate usually assumes.

When we start to search for sentient life on other planets, we will be taking with us a set of theoretical assumptions about what counts as life, and what counts as sentient life. When we encounter the crystalline space slugs of Ganymede, and observe the complex swirling movements they make when in proximity of each other, the Finean xenobiologists might conclude that since they do not grow, eat, or excrete, they are not alive. Furthermore, they might take this as evidence against other looser theories of what counts as life, because these space slugs (which count as alive according to the second theory) are clearly not alive according to our understanding of what life is. But the fact is that our understanding of life was based on a set of familiar data points: dogs, ants, plants, etc. Applying those opinions to this case is just not appropriate. To assume that the space slugs of Ganymede are not alive because they do not meet criteria established by an understanding that is based on an incomplete data set, and then to use this conclusion to argue against other theories of life, is to beg the question.

The Finean essentialist makes the same mistake as the Finean xenobiologist. In judging modal essentialism to be false because it is extensionally infelicitous with our understanding of essence, Fine assumes that our pre-philosophical opinions amount to such an understanding, that this understanding is extensive enough that it can be applied to resolve philosophical questions, and (most perilously) that it coincides with the neo-Aristotelian conception of essence. None of these assumptions are justified.

4. Antirealism

There are several things that one might mean by ‘antirealism’ about a certain realm of discourse. First, one might take an eliminativist approach and dismiss the discourse all together in the way Quine does for modality.¹² Second, one might aim to provide a reduction so that, whilst there are true claims to be made in the discourse, ultimately these are explained entirely in terms of some other feature of reality. Lewis takes such an approach to modality, reducing modal facts to non-modal facts about the way that worlds are.¹³ Third, one might claim that declarative sentences do not express propositions, instead favouring a non-cognitivist approach whereby declarative sentences do not have an assertoric function, but rather an imperative, directive, or emphatic function.¹⁴

The kind of antirealism I am concerned with in this project does not take any of these forms, but rather treats the subject matter of the discourse as not being true in virtue of the way the world is alone, but rather true in virtue of the way we represent the world. This is not to say that the subject matter is mind-dependent, any more than we might think that the location of the equator is mind-dependent. Rather, they rely on us in the sense that their truth conditions make indispensable reference to our relations to the world. This is in line with the kind of antirealism favoured by the likes of Blackburn, Dummett, and Wright.¹⁵

Ultimately, much of the debate can be elucidated with the often used metaphor of joint carving. On the one hand, you might see the job of the metaphysician as being similar to that of the butcher. When the butcher carves a carcass they do so at pre-existing joints that the butcher can readily identify. The process of carving is in no way relevant to the truth of claims about the joints. On the other hand, you might see the job of the metaphysician as being more like that of the cheesemonger. The

¹² Quine (1953)

¹³ Lewis (1986)

¹⁴ This is by no means an exhaustive list.

¹⁵ Blackburn (1984, 1993), Dummett (1959), Wright (1980, 1985, 1992)

cheesemonger cuts the wheel of cheese in accordance with no independent features of the cheese, but rather with their own interest relative priorities. Where the cheese is cut is not determined by any special feature of the cheese, but that is not to say that it is arbitrary where the cheesemonger cuts. Their choice will be affected by various factors, such as how much cheese they want in the portion, or what cuts are practical given the tools, etc. In just the same way, the antirealist about essence applies some kind of conventional process to the world (describing or thinking about it in a certain way, acting in a certain way) and that is the source of essentialist facts, rather than investigating the way the world is and discovering them. In accordance with this, essentialist facts may be the product of some kind of conventional process because thinking about the world in terms of essence is beneficial to us, or alternatively, essentialist facts may well be one of the tools that we use in order to gain some other kind of advantage.

Adopting an antirealist account of any discourse has the potential to render certain theoretical advantages. For example, such a theory may be more parsimonious, relying on fewer heavyweight metaphysical primitives, or it may allow us to better solve certain puzzles that occur in competing realist accounts. Such potential advantages provide some motivation for the present investigation. There is also good reason for such an investigation simply in virtue of the benefits that come with having a good idea of all of the theoretical options, what is sometimes called ‘mapping the logical space’. When we can examine all of the viable theories in a discourse against each other we are in a better position to draw more sophisticated conclusions. In this respect, investigations of essence are in something of a deficit.

5. Conventionalism

5.1 Traditional Conventionalism

Whilst it may be somewhat anachronistic to do so, I will call the view that the source of necessity is in our linguistic conventions *traditional conventionalism*. Prominent

conventionalists include Ayer, Carnap, and Wisdom.¹⁶ Much more needs to be said about what this means and how it relates to the larger project of antirealist essentialism. Talk of ‘the necessities’ or ‘the necessary truths’ is often ambiguous with regards to what particular claim is being talked about. On the one hand, ‘the necessities’ may mean claims of the form $\text{nec}(P)$, call this a *necessitation claim*. On the other hand it may mean the proposition P that is said to be necessarily true. Call this the *claim necessitated*.¹⁷ An example of a necessitation claim might be “it is necessary that no bachelors are married.” The corresponding claim necessitated in this instance is “no bachelors are married.”

Making this distinction allows us to be more exact in our specification of traditional conventionalism. We can say that a necessitation claim $\text{nec}(P)$ is true iff the claim necessitated P is true by convention. This establishes that the explanandum of the account is the truth of the necessitation claims. The explanans is the claim necessitated being true by convention. What does this mean? We can characterise the traditional conventionalist as claiming that the truth of a necessitation claim is guaranteed by linguistic conventions we hold about the meanings of words. What are conventions? According to Quine, conventions are the result of explicit agreements that stipulate rules governing a certain practice or activity, call these *decisive conventions*.¹⁸ Such conventions can take two forms; word-length conventions give terms their meanings, whereas sentence-length conventions stipulate the truth of one or more sentences based on some rule. For an example of a word-length convention, we hold that the term ‘bachelor’ is synonymous with ‘unmarried man’. Quine characterises conventionalism about the truths of logic in terms of sentence length conventions. He presents conventions as explicit stipulations that feature in a formal system (though not all elements of such a system need be decisive conventions, nor are all decisive conventionalist accounts required to set conventions within a formal system). According to Quine, conventions take the form of rules. For example “let any expression be true which yields a truth when put for ‘ q ’ in the result of putting a

¹⁶ Ayer (1936a, 1936b), Carnap (1937, 1947), Wisdom (1938)

¹⁷ One might similarly distinguish between the ‘claims analyticated’ and the ‘analytication claims’.

¹⁸ Quine (1936)

truth for 'p' in 'If p then q.'"¹⁹ As such, conventions require explicit decision (because they are the rules decided upon). Ayer on the other hand claims that conventions need not rest on explicit decision. He says:

Instead of saying that the laws of logic explicate meanings which have been deliberately assigned, can we not say that they explicate meanings which are implicitly adopted in current usage and do not require to be explicitly formulated in order to be understood?²⁰

Lewis takes a game theoretic approach to conventions, claiming that they are regularities in behaviour that instantiate *coordination equilibria*.²¹ A coordination equilibrium is a behavioural regularity formed in reaction to a coordination problem whereby non-deviation from this regularity results in mutual benefit, but where non-deviation from an alternative equilibrium could render equal benefit. For example, if we have the coordination problem of organising the roads in such a way as to avoid head on collisions, there are several options that one might take. Everyone could always drive on the right; everyone could always drive on the left; everyone could drive on the right in the mornings and the left in the afternoons, etc. The first two options are clearly equally as good as each other. There is nothing about driving on one side in particular that makes one option better, the important thing is that everyone always drives on the same side as each other. Note that the first two options are better than the third option (and other similarly gerrymandered options), as the third option is unnecessarily complicated. The first two options are coordination equilibria. It doesn't matter which we pick; they are the two best (and equally good) solutions. By adopting one, we adopt a convention. The important feature of this approach is that the conventions *are* the behavioural regularities, not the rules or stipulations that direct the regularities. As such, they need not be established by explicit agreement; instead they can be formulated retrospectively (or not at all). Call these *descriptive conventions*. Whilst descriptive conventions need not be the result of

¹⁹ Quine (1936), p.17

²⁰ Ayer (1936b), p.18

²¹ Lewis (1969)

explicit decision, they are not precluded from it. An example of an explicit descriptive convention might be the decision to drive on the left side of the road in the UK. In this case there was a coordination problem (the risk of head on collisions on the roads) and two available coordination equilibria that were just as good as each other (everyone driving on the right, or everyone driving on the left). Explicit decision determined that we chose one rather than the other, and so we came to have the convention of driving on the left side of the road. An example of an implicit descriptive convention might arise as follows. Two friends meet in the same pub every week. At first they specify a time and place, and reconfirm this for each subsequent meeting. When this has continued for a sufficient amount of time, if one says, “meet in the pub?” then it is implicit that this means the particular pub that they frequent. We can say that it is a descriptive convention that ‘the pub’, when uttered in this context, refers to one particular pub and no other. This is achieved without explicit decision.

5.2 The Downfall of Traditional Conventionalism

There are two stages to the attack on truth by convention. The first argues that convention cannot possibly account for the truth of all claims necessitated. This argument was most prominently made by Dummett, and Quine.²² The second argues the stronger claim that in fact convention is not in a position to account for the truth of any of the claims necessitated. This argument is repeated in many sources, including Hale, Cameron, and Sider.²³

Quine demonstrates that word-sized conventions, or definitions, are not capable of founding truths, but only of transforming them, saying that “what is loosely called a logical consequence of definition is therefore more exactly describable as logical truth definitionally abbreviated”.²⁴ As such, if there is to be any hope of accounting for the truth of the claims necessitated (and thus the necessitation claims, because for

²² Quine (1936), Dummett (1959). Quine (1951, 1960) also attacks analyticity.

²³ Cameron (2009, 2010a), Hale (2002, 2013), Sider (2003, 2011)

²⁴ Quine (1936), p.4-5

a necessitation claim to be true just is for the claim necessitated to be true by convention), then it must be through sentence-sized conventions that take the form of rules stipulating the truth of (infinitely large) classes of claims necessitated. These are the explicit conventions mentioned above.

Consider the truths of logic, which make up a portion of the claims necessitated. If not all of the logical truths can be accounted for, then *a fortiori* not all of the claims necessitated can be accounted for. Take T to be the base class of logical truths whose truth are stipulated by our conventions. T is infinitely large because it contains not only the results of direct single stipulations, but also general conventions that stipulate the truth of infinite classes of logical truths based on their form. However, T cannot contain all of the logical truths because there are infinitely many forms that logical truths can take and, finite creatures that we are, we can only hold finitely many conventions. As such, there will be sentences like s such that they are a truth of logic, but are not contained in T . If s is to be a claim necessitated then it must be a logical consequence of T . However, for every sentence s that is inferred in this way, there will be a sentence t , saying that $T \rightarrow s$. If we assume for the sake of argument that t is not a member of T then it must be a logical consequence of T . If this is the case then there must be a sentence u , saying that $T \rightarrow t$.²⁵ This sentence in turn needs to be explained in the same way as t . This results in an infinite regress of continually growing conditionals of the form $T \rightarrow (T \rightarrow s)$, $T \rightarrow (T \rightarrow (T \rightarrow s))$, etc. Any claim necessitated s that is not a member of T must be inferred from members of T , but this requires a logical rule of inference, and so the question must be asked again of this rule, and any rules required for it, *ad infinitum*. We encounter this problem because we already need to make use of logical rules in order to infer the validity of the inferences those logical rules license, from the conventions. This regress is vicious because we cannot make use of the logical rules we need without first helping ourselves to further rules which we are no more entitled to.²⁶ It seems that t cannot follow from T , but neither is it an option to claim that t is a member of T . There will

²⁵ u could also be written: $T \rightarrow (T \rightarrow s)$

²⁶ The regress is reminiscent of the Carroll regress (1895).

be an infinite number of sentences playing this role, and we would need an infinite number of individual stipulations, which is beyond our finite capacity for stipulation.

The second and more biting criticism of traditional conventionalism is the claim that conventions just aren't the kind of things that can make sentences true in any meaningful way. Just as definitions can only transform truth (by determining what a word means) and not found it, no convention can found truth, but rather can only determine the meaning of a sentence, that is, determine what proposition it expresses. The traditional conventionalist is accused of conflating that which determines meaning with that which determines truth. The claim is that whilst it is uncontroversial to claim that conventions determine which proposition a sentence expresses, whether that proposition is true is always a further matter beyond the conventions governing the sentence. This is something the traditional conventionalist does not recognise.

There are certain ways in which sentences can be true in virtue of the conventions we hold that are not sufficient to maintain a traditional conventionalist notion of truth by convention. First, a sentence that is about conventions may be true in virtue of those conventions. For example, there is a sense in which 'in the UK there is a convention that everyone drives on the left hand side of the road' is true in virtue of the conventions we hold, because it describes them accurately. It seems that it could also be the case that we might use our conventions to stipulate that a sentence is true, by stipulating that it expresses a proposition that is true. Neither of these senses of truth in virtue of convention is sufficient to maintain any kind of traditional conventionalist theory.

The conventionalist theories considered in this project do not rely on the troublesome truth by convention, so I shall not attempt any defence or extended discussion of it here. I take this to be enough to motivate the move from traditional convention, which I consider to rely on truth by convention, to what I call *neo-conventionalism*, a family of theories that maintain a role for convention, but steer away from truth by convention.

5.3 What Would a New Conventionalist Account Need to be Like?

For the purposes of this project I take the above considerations to be decisive against traditional conventionalism. As such, it will receive no more attention. However, there is a family of views, the neo-conventionalist theories, that whilst rejecting truth by convention, maintain a role for convention in securing the necessity of the claims necessitated. It is in using such theories that I provide an antirealist account of essence. Leaving aside the specifics of how these accounts reject truth by convention, they are still broadly conventionalist in nature, and as such they are held accountable to the same criteria that a good conventionalist theory must meet. Here I present these criteria, and for the rest of this project I take failure to meet them as a sign of inadequacy for any neo-conventionalist (or neo-conventionalist inspired) account.

A conventionalist account needs to specify what the account is for.

The philosophical goals of conventionalism are often phrased in neutral language such as the ‘explaining of’ or ‘accounting for’ necessities. It is not entirely clear however what it is about necessities that is being explained or accounted for. Conventionalism might be concerned with explaining how necessitation claims are made true, how we come to know them, or why we use modal discourse in the first place.

The conventionalism Quine describes is concerned with the source of the ‘necessitation claims’ truth. It is important to separate this from the distinct epistemic and pragmatic questions. An account might provide an answer to more than one of these questions. If an explanation of what makes the necessitation claims true is given in terms of the claims necessitated being analytic then how we come to know the truth of necessitation claims has also been explained. However, this may fall short of providing an answer to why we use modal discourse. Likewise, an account of why we use modal discourse in the first place may conflict with one’s ability to answer the other two questions. A non-cognitivist approach to modal sentences will preclude any account of what makes modal claims true, or how we come to know

them, because such claims have no descriptive content and thus are not truth apt or knowable.²⁷

I contend that in the context of conventionalism about necessity, a conventionalist's ability to address (or reject) all three of these questions is a distinct advantage. Indeed, it may even be a crippling deficiency of any proposed account if it cannot provide an adequate answer to them all (it is hard to see how an account that answers none could be an account at all). The primary philosophical goals of the accounts considered in by this project will be to account for the truth of essentialist claims. This has the consequence that it precludes non-cognitivist approaches, though such approaches are potentially available for the would-be conventionalist.

Note that this question is to be separated from the later question of what the conventions themselves are for. Whilst the question at hand focuses on what philosophers might have to gain by proposing such a theory, the question of what the conventions are for is concerned with the benefits for those who form the conventions in the first place. An alien anthropologist might form a conventional theory about why the humanoids on a certain green and pleasant landmass drive their automobiles on the left side of the road. The alien's theory explains why it is that we drive on the left side of the road; that is the benefit for them. What the conventions are for however, is the benefit that we, the convention formers, gain from those conventions. The person who proposed that we all drive on the left certainly wasn't looking to explain anything; they were looking to avoid collisions.

It needs to specify what the explanandum of the account is.

Just as the explanandum for the traditional conventionalist was the truth of the necessitation claims, the theories presented in this project must specify what the explanandum of their accounts are. The neo-conventionalist positions considered here face the additional constraint that if the explanandum is the truth of the necessitation claims then the explanans cannot be entirely a matter of convention,

²⁷ It is worth noting that whilst it does prevent these questions from being answered, it does not do so in a problematic way. There are no outstanding questions as they have, in effect, been rejected.

otherwise the account will fall victim to the criticisms of traditional conventionalism presented above.

It must specify whether conventions are implicit or explicit and how exactly the conventions achieve the desired result.

Once again, as above, the conventionalist must specify, at least to some extent, the nature of the conventions their account uses. This includes specifying if those conventions are explicit decisive conventions, or implicit descriptive conventions. There must also be an account of how those conventions secure the desired result. What are the actions that we, the conventional community, perform, and how do they achieve what the conventionalist account sets out to achieve?

Having explained the theoretical motivation for the theory, the account must specify what the practical motivation for the conventions themselves is. What practical advantage do they render for the conventional community?

This concerns what it is that the conventions are meant to achieve. For instance, when we form the convention to drive on the left hand side of the road, we do this so that we can drive on the roads, safe(er) from the risk of head on collision. The same applies with both decisive and descriptive conventions. When friends form the convention that when one says “meet in the pub?” the other knows that they meant a particular pub, this serves the purpose of allowing them to end up in the same place. Likewise, when we agree to licence a certain inference and never countenance its failure, there must be practical reasons that explain why that kind of inference is the kind of thing that it benefits us to make. Call this the ‘motivation’ of the account. Note that this is different from providing an explanandum of the account, or stating what philosophical purpose the account is for. In the cases of driving on the left, or meeting in the pub, neither the explanandum nor the philosophical purpose are what is being achieved by the convention. In the first case the explanandum is why we drive on the left side of the road. In the second case the explanandum is why ‘the

pub' refers to a particular pub as opposed to any other, or (heaven forbid) no pub at all. As such, regardless of whether a conventionalism about necessity is phrased in terms of necessitation claims or the claims necessitated, it must explain why we talk in terms of modality in the first place.

One might try to justify the account with the motivation of accounting for the necessities without having to refer to the real world, thus providing a simpler overall metaphysical theory. This cites parsimony as a theoretical advantage to motivate a conventional account. There are two kinds of advantage that a conventional system might yield. The first is theoretical advantage of the kind just mentioned; this benefits those who theorise about the topic in question. The second is a practical advantage to those who were party to the formation of the convention (and those who continue to benefit from its continued use). To present the motivation as the former kind of advantage is to conflate the two. The motivation must be pre-theoretic advantage of the latter kind. It must be a motivation that benefits those who first adopted it. Remember that descriptive conventions originate in coordination problems that require a solution. There are no (obviously relevant) coordination problems that hang on matters of theoretical advantage (even if there were, they could not arise before the implementation of the convention inside the field they study). Conventional accounts (both implicit and explicit) must be motivated by practical advantage.

In as far as the motivation of the conventions are concerned, it is worth noting that the motivation may not be even related to the explanandum of the account. For instance, adopting modal language may merely be the means by which some other ends is achieved. The product, as it were, of the conventionalist theory, may even be an unintentional by-product of achieving an apparently unrelated end.

5.4 Conventionalism vs. Antirealism

It is important to note that, whilst the antirealist account of essence to be developed here uses conventionalist mechanisms, it is not strictly speaking a conventionalist

account, as it differs in at least one important respect. It is, I take it, an essential feature of conventions that they are, to some degree, a matter of choice. The coordination equilibria that Lewis takes conventions to represent require that there be an equally viable alternative.²⁸ When we decided to drive on the left hand side of the road, we could just have easily decided that we drive on the right. Which side we drive on is not the matter of importance. What is important is that we all drive uniformly on the same side, and that we all believe that we all drive uniformly on the same side (and that we all believe that we all believe this, and so on); only then will we avoid constant head on collisions.

The same goes for all conventions. However, I do not take this to be an essential feature of the antirealist essentialism I develop in this project. Here I distinguish between that which is *strictly* conventional (in the Lewisian sense), and that which is *broadly* conventional.²⁹ It is essential to a strict convention that there should be an available alternative convention that we might have held that would have been equally good for the job. A broad convention, however, differs in that (and only in that) it does not require an available alternative.

It is reasonable to think that there could be situations entirely like those involved in strictly conventional cases but in which we could not have chosen otherwise (for instance, where there are multiple coordination equilibria, but we are unable to choose one or more of them). Consider the example of us driving on the left hand side of the road in the UK. Imagine that, through some evolutionary quirk, when confronted with a choice between left and right, humans have a strong predisposition to always choose the left. Whilst driving on the right would have been just as good, we were never going to pick it. It was inevitable that we chose to drive on the left, but this still qualifies as a case of broad convention. Likewise, if the person/people who invented chess had been cognitively constricted in such a way that the board had to be the size it was, this is no less a case of broad convention.³⁰

²⁸ Lewis (1969)

²⁹ Sidelle (1989) uses convention in an even broader sense, whereby ‘convention’ simply means any mind-dependent contribution.

³⁰ These examples are not intended to motivate the claim, merely to clarify it.

The world may have limited our choices (perhaps all the way to a single choice), but this does not compel us to adopt a realist perspective on the rules of chess, or the laws governing how we drive. Nor does it mean that we must adopt a significantly different form of antirealism. I use ‘broad convention’ to describe the kind of mechanism that is similar to strict convention, but where it is not clear that we could have done otherwise (for example, if our brains were hardwired in such a way as to make us choose one option). Broad convention is still established through the same kind of coordination problem set up that strict convention uses. It merely does not require that we be able to choose an alternative.

Whilst conventions and the conventionalist method feature heavily in this project, few, if any, of what I will call conventions must actually be conventions in the strict sense stated here. I continue to call them conventions out of convenience, though I will not call the final position ‘conventionalism’. Instead, I call the theory of essence developed in Chapter Four an antirealist theory of essence that uses a conventional mechanism.

6. Chapter Outline

In Chapter One I present a neo-conventionalist theory of modal essence based on the neo-conventionalist theories of modality adopted by Ross Cameron and Theodore Sider.³¹ I explain the advantages that neo-conventionalist theories have over traditional conventionalism, and investigate exactly what role convention plays. I defend neo-conventionalism from the contingency problem, the claim that the contingency of our conventions jeopardises our ability to maintain the characteristic S4 principle because the source of necessity is contingent. In defending against this problem I look at responses from Crispin Wright and Iris Einheuser, concluding that whilst on the right track, they need modification to be completely satisfactory.³² Whilst I conclude that neo-conventionalism is safe from the contingency problem, the positions presented by Cameron and Sider are not satisfactory because they fail to

³¹ Cameron (2009, 2010a, 2010b), Sider (2003, 2011)

³² Einheuser (2006), Wright (1980, 1985)

adequately meet the criteria for good conventionalist theories laid out in Section Five of this introduction.

In Chapter Two I examine another position that I consider to be a neo-conventionalist theory, presented by Alan Sidelle.³³ Using Sidelle's neo-conventionalism about modality I give an antirealist account of modal essence similar to that given in Chapter One, but with important developments that leave it able to adequately match the criteria imposed in Section Five of this introduction. I address some potential problems that come with Sidelle's account. These include Gillian Russell's indexicality problem, which claims that because necessity is a property of propositions, and analyticity is a property of sentences, we should not expect linguistic accounts of necessity to work. Another problem is that Sidelle's account assumes the viability of, and makes use of analyticity. As this project works on the assumption that analyticity is to be rejected, this is a problem. I respond by arguing that using the tools that Sidelle allows himself, one can dispense with analyticity and proceed without it. I conclude that the position developed sufficiently demonstrates that antirealist modal essentialism is a viable position to hold, if modal essentialism is a viable position. In the next chapter I address this important caveat.

In Chapter Three I present neo-Aristotelian essentialism as advocated by Kit Fine.³⁴ The work I do in this chapter comprises two tasks. The first task is giving an exposition of Fine's essentialism, and making certain modifications relating to the way Fine develops his account. The second task is considering Fine's now famous arguments against modal essentialism. Whether the position developed in Chapter Two is truly a viable position depends on whether modal essentialism in general is viable; Fine argues it is not. I argue that, based on the interpretation of the debate I gave in Section Three of this introduction, we should not take Fine's arguments to be decisive against the modal essentialist. Instead of taking the arguments as showing that modal essentialism is extensionally infelicitous with a well-defined understanding of essence furnished by our pre-philosophical opinions, we should

³³ Sidelle (1989)

³⁴ Fine (1994, 1995a, 1995b, 1995c, 2000, 2007)

take them as better defining the proposed understanding of essence that Fine advocates and how it differs from modal essentialism. In addition, I present a form of realist modal essentialism that is not subject to Fine's objections. I conclude that, whilst Fine's arguments against modal essentialism are not decisive, neo-Aristotelian essentialism is the more motivated position. This is because under the neo-Aristotelian interpretation, essence is much more interesting and useful. Neo-Aristotelian essence fits into our wider metaphysics. Knowing the essentialist facts tells you something new that you cannot know just by knowing other things, and Fine claims that metaphysical modality can be reduced to essence. Contrast this with modal essentialism, where knowing the essentialist facts doesn't tell you anything new above the modal facts, and essence doesn't play much of a theoretical role in our wider metaphysics (except in so far as the modal facts do). As such, whilst modal essentialism is a viable position, neo-Aristotelian essentialism is better motivated. With this in mind, I turn to providing an antirealist account of neo-Aristotelian essence.

In Chapter Four I employ the mechanism from Sidelle's conventionalism to develop an antirealist account of neo-Aristotelian essence. This is a more complicated task than using Sidelle's work to develop an antirealist modal essentialism, and I consider the various forms that such an account might take. In particular, I investigate whether such an account must commit one to antirealism about objects, as Sidelle insists. I conclude that it need not, but that the position where one is an antirealist about objects is more theoretically elegant. Finally I take some steps to show that there are some advantages that antirealist neo-Aristotelian essentialism has over Fine's realist counterpart.

In Chapter Five I consider essentialist theories of necessity. One of the primary advantages of neo-Aristotelian essentialism over modal essentialism is that it has the potential to provide an account of modality. Here I explore how such an account might go, looking at examples from Fine, Fabrice Correia, and Bob Hale.³⁵ I conclude that whilst all of the realist options are viable, none are without their own

³⁵ Fine (1994, 1995c), Correia (2006, 2012), Hale (2013)

challenges. I then present an antirealist essentialist theory of modality along similar lines, and argue that it is better suited to responding to these challenges than its realist competitors.

The project sums up with some remarks about the antirealist theory of essence that has been developed. I conclude that antirealism about essence is a worthwhile position that warrants further discussion in contemporary debate, and I highlight some specific areas of further work to be done.

Chapter One: Deflationary and Humean Neo-Conventionalism

1. Introduction

Following the widespread repudiation of truth by convention detailed in the introduction, a family of contemporary theories of modality have arisen that retain a role for convention, whilst retreating from the claim that any propositions are true by convention. I call these positions *neo-conventionalisms*. Notable proponents of neo-conventionalism include Ross Cameron, Alan Sidelle, and Theodore Sider.³⁶ In this chapter I examine Cameron's deflationary account and Sider's Humean account. I examine these accounts as two similar forms of neo-conventionalism, and present them as a means by which we can achieve an antirealist account of essence. By providing an antirealist account of modality, neo-conventionalism furnishes us with a strategy for antirealism about essence as construed by the modal essentialist. We can claim that the essentialist facts are grounded in the modal facts, but that the modal facts are in turn conventional in nature.

In Section Two I give an exposition of the deflationary and Humean accounts. I present them together as I take them to be two ways of making the neo-conventionalist move away from truth by convention. I discuss the use of naturalness in the deflationary position, and motivate discussing the position in terms of convention rather than unnaturalness. In Section Three I give the claim that neo-conventionalism does not rely on truth by convention a closer examination. I give a more fine-grained account of the roles convention might play, and conclude that whilst one might characterise neo-conventionalism as relying on a form of truth by convention, it does not do so in any way that is problematic. In Section Four I consider the contingency problem, the claim that the contingency of our conventions

³⁶ Cameron (2008, 2009, 2010a, 2010b), Sidelle (1989, 2009, 2010), Sider (2008, 2011)

jeopardises their ability to play a role in grounding necessity. I examine a response from Iris Einheuser, suggest how it might be improved, and show how it can be applied to the deflationary and Humean positions. In Section Five I conclude that two significant deficiencies of the deflationary and Humean positions render them inadequate without significant work, work that Sidelle's neo-conventionalism, the subject of Chapter Two, engages in.

2. Neo-Conventionalism

2.1 Deflationary and Humean Neo-Conventionalism

In the Introduction we distinguished between necessitation claims and the claims necessitated. A necessitation claim is a claim of the form 'necessarily P' that is an attribution of necessity. A claim necessitated is a claim 'P' to which necessity is attributed (but need not be an attribution of necessity itself). For the necessitation claim 'necessarily P', the claim necessitated is P. The traditional conventionalist can be characterised as claiming that the claims necessitated are true by convention and that this is what makes the corresponding necessitation claims true. Neo-conventionalism retreats from this position. It still attributes the necessity of the claims necessitated to convention, but rejects truth by convention. One might think that by retaining some role for convention the neo-conventionalist is moving from truth by convention for the claims necessitated to truth by convention for the necessitation claims. This however is a conclusion that the neo-conventionalist resists. Neither 'vixens are foxes', nor 'necessarily, vixens are foxes' is true by convention. Neo-conventionalist explanations of necessitation claims like 'necessarily, vixens are foxes' take the form: 'necessarily, vixens are foxes' is true iff some fact F holds, where what counts as an instance of F is determined by convention. As such, the necessitation claim is true not in virtue of convention, but in virtue of F holding.

Here I consider two forms of neo-conventionalism; Cameron's deflationism, and Sider's Humean account. Cameron's position starts with an actualist possible worlds

approach to modality. For P to be necessitated is just for P to be true in all of the possible worlds; for P to be impossible is for not-P to be necessitated, that is, for P to be true in none of the possible worlds. Likewise, for P to be possible is for it to be true in at least one world; for P to be contingent it must also be false in at least one world. Worlds are abstract entities (Cameron's preference is for sets of propositions). However, such accounts are often faced with challenges to their reductive ambitions. They have a problem explaining why some worlds are possible and others are impossible. This allegedly cannot be done without taking some modal notion (often consistency) as primitive. What distinguishes the possible worlds from the impossible ones? Cameron claims that there is such a distinction, but that it is an *unnatural* one. To borrow the metaphor, the distinction does not carve reality at nature's joints; making a distinction that is not reflected in the deep metaphysics. The distinction between the possible and impossible worlds is genuine and mind-independent, but highly unnatural.³⁷ That is, there is a distinction, and it doesn't depend at all on what we think about it, but it was decided upon by us, and does not represent any metaphysical difference between the things distinguished (for instance, a natural property shared by all of one group and none of the other). All that distinguishes between members of one group and members of the other is that some are assigned to one group, and some to the other. The distinction is real, it's just nothing special. Frege gives similar remarks about the equator:

We often speak of the equator as an *imaginary* line; but it would be wrong to call it a *fictitious* line, it is not a creature of thought, a psychological process, but is only recognized or apprehended by thought. If to be recognized were to be created, then we should be able to say nothing positive about the equator for any period earlier than the date of its alleged creation.³⁸

To use a new analogy: I have invented a new board game, and I'm deciding on the size of the board. To help me decide I have a 100cm by 100cm square board in front

³⁷ Cameron (2009), p.13

³⁸ Frege (1884), p.35

of me that is divided into 10,000 1cm^2 squares. There are many potential sizes I can pick: 10x10, 30x30, 50x50, 30x50, etc. I pick 30x30 because of reasons relevant to the game mechanics, ergonomics, etc., but the other options were also viable (in particular, at least some of the alternatives were viable considering those same considerations relevant to the game mechanics). Once I make it, the board is genuinely that size. It is not mind-dependent that the board has a 30x30 layout, nor that it has a 900cm^2 surface area. However, it is a matter of convention that the game board is that size, because I could just as easily have picked another size.³⁹ This does not mean that there had to be nothing in the choice. I had good reason to pick that size; it just wasn't a reason that had anything to do with metaphysics.

Deflationism aims to explain the truth of the necessitation claims. Questions concerning the epistemology of modality, or why we modalise are left unanswered. The analysis states that a necessitation claim $\text{nec}(P)$ is true iff the claim necessitated P is true in all of the possible worlds. The role of convention in the account is to ground the difference between the possible and impossible worlds. These conventions are presumably implicit, descriptive conventions, though little is said about them.⁴⁰

Sider presents an account which is similar to the deflationist account, and can be presented without reliance on possible worlds.⁴¹ Instead of claiming that $\text{nec}(P)$ iff P is true at all possible worlds, Sider says that $\text{nec}(P)$ iff P is true and P is a claim of a certain kind. The kinds of claim in question are selected by convention (in the same way the deflationist claims that the possible worlds are selected by convention). For example, if analytic, logical, and mathematical claims were the kinds selected, then it would be true that $\text{nec}(P)$ iff P is a true claim that is analytic, logical, or mathematical. In effect, ' $\text{nec}(P)$ ' means ' P is an analytic, logical, or mathematical claim'.^{42, 43}

³⁹ In Section Three I take a closer look at exactly what role convention in playing.

⁴⁰ Indeed, Cameron doesn't explicitly mention conventions at all.

⁴¹ Sider (2003, 2011)

⁴² Sider (2003), p.28

⁴³ Sider is here working with an idiosyncratic notion of analyticity. For Sider, a claim is analytically true when it is true, and 'a definitional statement'. In this context, the truth of the claim has nothing to do with the fact that it is a definitional claim. As such, analytic propositions are not instances of truth by convention.

Convention selects a set of true sentences to act as modal axioms; then selects a set of truth preserving relations between groups of sentences and sentences to act as modal rules.⁴⁴ For any such collection there will be a corresponding set of modal theorems (the closure set of the modal axioms under the modal rules, which is the intersection of sets that contain every modal axiom, and contain every sentence to which a subset of its members bears a modal rule relation). According to the Humean, these modal theorems are the claims necessitated. This framework gives the Humean the ability to accommodate necessity in any realm of discourse. In order for the theory to account for some branch of necessity, it need only add the appropriate claims to the modal axioms.

2.2 A Note on Naturalness

Cameron's deflationary account makes significant use of the natural/unnatural distinction. Use of this notion seems to be at odds with the spirit of the project, encouraging the adoption of unnecessarily realist metaphysics. Here I consider the distinction, and argue that deflationism is better thought of in terms of conventions than the natural/unnatural distinction. Cameron claims that the distinction between the possible and impossible worlds is unnatural. Talk of naturalness goes back to Lewis, but is a less than clear notion. It is often explained with the equally opaque metaphor that natural distinctions 'carve reality at its joints', whereas unnatural ones do not. Cameron presents naturalness in terms of objective similarity. If there is a natural distinction between the Fs and the non-Fs then there is a level of objective similarity among the Fs. However, if the distinction is an unnatural one, then there needn't be such similarity. According to Lewis, talk of properties alone does nothing to capture facts about resemblance. In his words "[p]roperties carve reality at the joints – and everywhere else as well."⁴⁵ To capture resemblance facts, we need to single out a special class of properties, the natural properties. The sharing of natural properties makes for resemblance among the things that share them. This distinction

⁴⁴ Sider (2011), p.320

⁴⁵ Lewis (1983), p.346

is objective, and according to Lewis preferably one of degrees. There are *perfectly natural* properties, classes of objects that all and only instantiate one specific universal; *less natural* (or *imperfectly natural*) properties that share in “families of suitable related universals”;⁴⁶ finally there are *unnatural* properties, whose members presumably fail to satisfy either criterion. An example of a perfectly natural property is having charge +1. An example of an imperfectly natural property is being metallic; although its members share no one universal of ‘being metallic’, there is a family of universals (those that we associate with the properties common to metallic objects) that they do share. An example of an unnatural property is Goodman’s *grue*.⁴⁷

A property’s being natural relies on there being a corresponding universal (or group of universals). A property like *grue* is (allegedly) unnatural because, whilst all of its members share a property (that of ‘being *grue*’), they need not share any universals. This raises an epistemic question. How do we know which properties have corresponding (groups of) universals? There is no obvious answer to this question, and without such an answer we cannot know which properties are natural. Likewise, it begs the question to assume that the green things share a kind of resemblance that the *grue* things do not. After all, if there were universals for *grue* and *bleen*, and not for green and blue, then *grue* (and not green) would be perfectly natural, and we would be in an epistemically identical situation.

Furthermore, whilst *grue* is often taken as a quintessential unnatural property, it’s not clear that it is under Lewis’ understanding. *Grue*’s alleged unnaturalness is based on the assumption that green and blue have corresponding universals, and that *grue* does not. Assuming that this is correct, and assuming that we can know this, on the above understanding of naturalness *grue* should be classified as natural. When one considers the set of all *grue* things, there are two universals that they all share (considered atemporally): green and blue. Considered at any time *t* they share one universal (either green or blue, depending on when *t* is). This is not just coincidence.

⁴⁶ Though quite how we are supposed to know that/how certain universals are related to each other is not clear.

⁴⁷ Goodman (1983)

For any time t and any two possible grue things, if they fail to share one of those two universal then they cannot both be grue.

Further still, such an understanding of naturalness requires an ontology of both properties (as sets of objects) and universals. However, one might think that the contemporary notion of naturalness, the one that is currently under consideration, need only gain its intuitive force from Lewis' account. As such, one need not incur the ontological costs, taking naturalness as primitive.⁴⁸ I'm not convinced that this would be a legitimate move. Lewis provides an account that explains exactly what we mean when we say a property is natural; without the elements involved in the account it is just not clear that the intuitive notion survives (at least in a form any more resilient than the intuitive one we started out with). There certainly is an intuitive notion available however. When one thinks about chemical kinds, such as gold, there is more that the instances of gold have in common than the instances of 'yellow metal' do. However, the lesson that grue teaches us is that our intuitions don't provide us with metaphysical entitlement in these cases. Imagine that there were only two yellow metals, gold, and fool's gold. We might then give 'yellow metal' the disjunctive definition 'being gold or fool's gold' (alternatively, if we take properties as sets, the union of the sets $(G \cup F)$ of gold objects and fool's gold objects respectively). However, this is based on the assumption that gold and fool's gold are natural kinds, and that yellow metal is not. To work on this assumption is to beg the question. If one were to assume that yellow metal (and not gold or fool's gold) was a natural kind, then one might give 'gold' the conjunctive definition 'being a yellow metal and (for instance) being valuable' (or, the intersection of the sets $(Y \cap V)$ of yellow metal things and valuable things, respectively).

Taking talk of naturalness seriously potentially requires one to invest deeply in realist metaphysics. Even if one is willing to do this, the epistemic problems remain. It seems problematic to hold naturalness as anything more than an intuitive notion such as "a property, kind, notion, or distinction is natural if application is

⁴⁸ Indeed, despite his comments in his (1983), Lewis is often reported to have taken naturalness as primitive.

guaranteed/precluded based on a single specific metaphysical difference”.⁴⁹ Here I will avoid talking in terms of naturalness (except when using the intuitive notion). I will talk of the difference between the possible and impossible worlds as conventional. Once we commit to thinking about the distinction as conventional (not relying on talk of naturalness for anything more than heuristic purposes) it is important that we establish exactly what we mean by this claim. The role that convention plays is discussed in Section Three.

3. Truth by Convention

3.1 Does Neo-Conventionalism Rely on Truth by Convention?

The neo-conventionalist move is an explicit and complete retreat from truth by convention for the claims necessitated. Instead, convention plays a role in the necessity of the claims necessitated. It would, however, be a mistake to think that this involves the truth by convention of the necessitation claims. The necessitation claims are true because the claims necessitated are true at all of the possible worlds; they are not true by convention. The role of convention is merely to determine which worlds are possible. To borrow from his metaphor, Cameron is Scottish because he was born in Glasgow, not because of the conventions that drew the borders of Scotland so as to include Glasgow. Though convention does play some role in the truth of $\text{nec}(P)$, it is a much smaller role than the traditional conventionalist had in mind.

So what exactly is the role of convention in the truth of necessitation claims? There are three identifiable roles convention might play. For lack of a better name I will call these conventional grounding types 1, 2, and 3. A sentence P is type 1 grounded in convention when P is true iff some convention C holds (where C does not feature in the meaning of P , for instance if P were an assertion that we hold convention C). P is type 2 grounded in convention when P is true iff some other fact

⁴⁹ Quite how imperfect naturalness is to be understood without circularity, or introducing some new notion practically synonymous with naturalness to account for what constitutes a family of suitable related universal, requires a substantial research project in its own right, and is well beyond the scope of this project. I prefer to conveniently side-step the issue by introducing convention instead.

F holds, where C determines that the truth of P is contingent upon F.⁵⁰ P is type 3 grounded in convention when P is true iff some fact F holds, where what counts as an instance of F is determined by C.⁵¹

Cameron presents an account whereby the truth of sentences of the form *nec*(P) is not type 1 grounded in convention, but is type 3 grounded in convention. *Nec*(P) is true because P is true in all of the possible worlds; which worlds are possible is determined by convention. Cameron does not comment on whether the truth of *nec*(P) is type 2 grounded in convention. Here I answer three questions. First, is it conventionally determined that *nec*(P) is true iff P is true in all of the possible worlds (that is, is *nec*(P) type 2 grounded in convention)? Second, if *nec*(P) is type 2 grounded in convention, is that sufficient for it to count as true by convention? Third, if it is an instance of truth by convention, is this problematic for the deflationist?

Even if the distinction between possible and impossible worlds remains unnatural, it seems that there is some level of underlying metaphysical modality. There are still truth-at-world facts; we can still talk about actual and counterfactual situations. It seems plausible that there may be a genuinely modal landscape independent of how we carve it into the realms of possibility and impossibility, just as there was a geographical landscape independent of how we distinguished between England and Scotland. Even if this were the case, without the distinction between the possible and impossible worlds there can be no non-conventional theory of necessity. Ought we conclude from this that not only is it conventional which are the possible worlds, it is also conventional that “*nec*(P) iff P is true at every possible world”? Yes. It is highly implausible (though admittedly not contradictory) that it should be a metaphysical fact that *nec*(P) iff P at every possible world, without the distinction

⁵⁰ It is worth noting that type 2 grounding in convention can furnish us with an alternative strategy for a conventionalist account of modal essence. Instead of taking the strategy of this chapter by analysing essence in terms of modality and then be a conventionalist about modality, the alternative is to have some non-conventional (perhaps even realist) theory of modality, and then claim that essentialist facts are conventionally grounded in modal facts. That is, by claiming that essential facts are type 2 conventionally grounded in modal facts. This is not an option that I shall pursue here.

⁵¹ Types 2 and 3 might be better described as cases of partial grounding.

between possible and impossible worlds also being metaphysical.⁵² God wouldn't decide that necessity depends on truth at all possible worlds without deciding which of the worlds are possible (or doing anything in the metaphysics to distinguish between the two).

It is interesting to note that this need not always be the case. Imagine I'm a member of a cult, and I'm interested in making prudent decisions. As a representative of the cult, whether a course of action is prudent is determined by whether it is in accordance with the cult's dogma. The dogma of the cult is a list of goals conventionally determined by its leaders. Some action A that I'm considering will be prudent (or: the proposition <A is prudent (for me now, as a representative of the cult)> is true) iff A is in accordance with the dogma of the cult. This is a case of type 3 grounding (because it is conventionally determined what counts as being in accordance with the cult's dogma), but not a case of type 2 grounding (because what makes an action prudent need not be conventional). Likewise, I'm a factory worker in a communist country, and I subscribe to the goals of communism. Therefore it is prudent for me to make whatever it is that the factory managers tell me to make (so that I can do my part in serving the greater good/avoid the harsh penalties of my masters). Little do I know, the managers determine what the factory produces on a whim. Their goal is to keep me and my comrades working so that we won't cause trouble, not to produce anything of any worth. In this case, the prudence of, or perhaps even the morality of, my actions are type 3 grounded in convention, but need not be type 2 grounded.

The reasoning in this case seems to go like this. The A facts (facts about whether given actions are prudent or not) are non-conventionally grounded in the F facts (consistency-with-doctrine facts). The As are divided exclusively and exhaustively into the Bs and Cs (prudence and imprudence facts); likewise the Fs are divided into the Gs and Hs (consistence and inconsistency-with-doctrine facts). The Bs are type 3

⁵² So if 'nec(P) iff P at every possible world' being a metaphysical fact entails that the distinction between possible and impossible worlds is a metaphysical one, and that distinction is not a metaphysical one, then by modus tollens we should conclude that the biconditional is non-metaphysical also.

grounded in the Gs and the Cs are type 3 grounded in the Hs. The distinction between the Gs and Hs is an unnatural one; whether an F is a G or an H is a matter of convention (because whether an F is a G or an H depends on what actions are condoned by the doctrine, which is conventionally selected). That a given A is grounded in an F is not conventional (it is not conventional that acting in accordance with the doctrine is prudent), as is the grounding of given Bs and Cs in Gs and Hs respectively (because the prudence of an individual action is non-conventionally determined by accordance with the doctrine). The difference between As and non-As, and Fs and non-Fs is natural, but the unnatural sub-division within the Fs make the grounding of As in Gs type 3. There's no reason to think that there is also type 2 grounding occurring.

One can attempt to draw a modal analogue of such a case. Modal facts are non-conventionally grounded in (sets of) truth-at-world facts.⁵³ The modal facts are divided into those that are consistent with our standard modal notions (possibility, impossibility, necessity, and contingency), and those that are not (schmeccessity, schmimpossibility, and schmossibility).

The set of modal facts (M) is exhaustively divided into two proper subsets: the set of modal facts that feature our standard modal notions (M_s), and the set of modal facts that feature non-standard modal notions (M_{-s}). The set of truth-at-world facts (W) is exhaustively divided into two proper subsets: the set of those truth-at-world facts that only involve possible worlds (W_p), and the set of truth-at-world facts that involve one or more impossible worlds (W_{-p}).

The modal facts that feature our standard modal notions (M_s) are type 3 grounded in various subsets of the power set of the set of truth-at-world facts that involve only possible worlds; the modal facts that don't feature our standard modal notions (M_{-s}) are type 3 grounded in subsets of the power set of the truth-at-world facts that involve one or more impossible worlds. The distinction between the

⁵³ Here I use 'modal facts' in the particularly broad sense that any fact we would express using a modal operator is a modal fact. This includes facts that would need to be expressed by modal notions we do not possess. For example, schmeccessarily(P), and schmossibly(Q), as well as necessarily(R), and Possibly(S) all express facts that feature in the set of modal facts.

possible and impossible worlds is a matter of convention, as is the distinction between the modal facts that feature our standard modal notions and those that don't. In this scenario there is a set of modal facts M , the members of which are grounded in various (sets of) members of W , the set of truth-at-world facts. The conventions select subsets of these sets as being of interest, but that the individual facts are grounded in each other is non-conventional. M_S facts are type 3 grounded in (sets of) W_P facts (the set of truth-at-world facts that involve only possible worlds). Are they type 2 grounded? Yes. Whilst it is non-conventional that each member of M_S is grounded in whichever W_P facts they are grounded in, it is conventional that our modality is the one featuring the notions of M_S . The link between our modal notions and truth-at-world facts is also conventional. We assign a possible world semantics to our modal notions, theorising that possible worlds non-conventionally ground modal facts. This is done by convention.

The reason why the conclusion differs in the modal case from the earlier cases is because the distinction among the facts being grounded is an unnatural one. This results in the grounding being of type 2 as well as type 3. Admittedly, this is not a demonstration that there cannot be a modal case of type 3 grounding without type 2 grounding, but I hope that by showing this candidate (which I take to be the best candidate) to also be type 2, the reader will have sympathy with the claim. It appears that if modal facts are type 3 grounded in convention, then they must be type 2 grounded in convention as well.

3.2 Does Type 2 Grounding Count as Truth by Convention?

The next pressing question is whether cases of type 2 grounding count as cases of truth by convention. One might think that when 'P iff F (where the contingency of P upon F is determined by convention C)' is the case, P is true by convention. The truth of P may not be 'grounded in convention', but one might describe the truth of P as 'conventionally grounded'. For this to be the case (at least in any interesting sense) it is important that the role convention is playing is not merely one of assigning

meaning to P. “There is a spoon on the table” is true iff there is a spoon on the table. That the truth of that sentence is contingent upon the presence of a spoon is determined by convention, but only in the uninteresting sense that the convention assigns the meaning to the sentence of the proposition <there is a spoon on the table>. If type 2 grounding is going to be of interest then it had better not be of this form.

There are examples of type 2 grounding that go beyond the assignment of meaning. That two people are married is type 2 grounded in them having a marriage ceremony of some kind. This is clearly conventional, yet we don’t think that the meaning of ‘being married’ has anything to do with them having had a ceremony. When I say ‘x and y are married’ this means that they bear a certain relation to each other, not that they both partook in a ceremony. Likewise for declarative utterances; the ship is called ‘Matilda’ because the Queen broke a champagne bottle on it. Saying ‘this ship is called Matilda’ means that the name of the ship is ‘Matilda’. It does not mean that Queen broke a champagne bottle on it (even if that is a requirement for the ship to have that name). Is there reason to think that necessitation claims can be type 2 grounded in this way? Yes. It seems that if the necessitation claims are type 2 grounded in convention, then this is sufficient for us to say that they are true by convention. Consider the Quinean interpretation of how a reductive conventionalism would look. It would contain explicit conventions of the form:

(II) *Let any expression be true which yields a truth when put for ‘q’ in the result of putting a truth for ‘p’ in ‘if p then q.’*⁵⁴

Such conventions stipulate the truth of sentences that fulfil certain criteria. A reduction for the truth of R might appear in the disjunctive form of all such conventions:

R is true iff R yields a truth when put for ‘Q’ in the result of putting a truth for ‘P’ in ‘if P then Q’, or yields a truth when put for ‘Q’ in the result of putting a falsehood for ‘P’ in ‘P or Q’, or... etc.

⁵⁴ Quine (1936), p.17 (Quine’s italics)

According to this reduction R is true iff it satisfies certain conditions. The reduction itself contains no explicit mention of conventions at all, but it is type 2 conventionally grounded that these are the conditions that R's truth is contingent upon. This, if anything, is taken to be a classic example of the notion of truth by convention. If we're not willing to count this as a case of truth by convention then we're simply working with a different notion than the traditional conventionalists were (and we may well be). This isn't necessarily a bad thing. If we don't want to call this truth by convention then that's fine, but to then repudiate a notion after excluding the instances of it that are (allegedly) plausible isn't fair. It's moving the goalposts.

Ultimately it's immaterial what we call it. I doubt the critics of truth by convention would be concerned about whether what they're attacking should properly be called truth by convention. The notion they're attacking is indefensible, but that's nothing new. Quine taught us that conventions determining meaning could only transform truths, not found them.⁵⁵ The notion of truth by convention that I'm using is broader than that. I will proceed on this notion, but the reader is welcome to substitute another name into its place at their leisure.

3.3 Is Reliance on Truth by Convention Problematic for the Deflationist?

Gillian Russell presents four kinds of determination in her examination of the role meaning plays in the truth of sentences.⁵⁶ These kinds of determination are: partial, conjoint, full, and redundant. Establishing just how convention determines the truth of the necessitation claims will reveal whether or not it is the problematic kind of truth by convention used by the positivists and rejected by the neo-conventionalist.

We can understand functions as sets of ordered n-tuples of arguments and their outputs. For example, the addition function is the set of n-tuples $\langle x, y, z \rangle$, where z is the output of the addition of x and y . In language there is a function from the meaning of a sentence (the proposition it expresses) and the way the world is to the truth value of that sentence. This function is the set of triples $\langle m, w, v \rangle$, where m is

⁵⁵ Quine (1936), p.4-5

⁵⁶ Russell (2008), p.35

the meaning of the sentence, w is the corresponding (part of the) world, and v is the truth value. The role that any part of those tuples plays can be explained by the kind of determination it does. Relative to a function F , the four kinds of determination can be defined as follows:

Partial determination: an argument place i partially determines result y of F iff there exist two $n+1$ -tuples in F $\langle x_1, \dots, x_i, \dots, x_n, y \rangle$ and $\langle x_1', \dots, x_i', \dots, x_n', y' \rangle$ such that their arguments $x_1 - x_n$ and $x_1' - x_n'$ are identical save for x_i and x_i' , and they differ in their result ($y \neq y'$). That is, difference in i alone can change the result.

Conjoint determination: a series of argument places $i - k$ conjointly determine result y of F iff there exist no $n+1$ -tuples in F $\langle x_1, \dots, x_i, \dots, x_k, \dots, x_n, y \rangle$ and $\langle x_1', \dots, x_i', \dots, x_k', \dots, x_n', y' \rangle$ such that the arguments $x_i - x_k$ and $x_i' - x_k'$ are identical, but they differ in their result ($y \neq y'$). That is, constancy of the values of $i - k$ guarantees the result.

Full determination: a series of arguments $x_i - x_k$ conjointly determine result y of F iff for every member of F , if positions $i - k$ are $x_i - x_k$ then the result is y . That is, the presence of $x_i - x_k$ guarantees y . (Note that this differs from conjoint determination in that full determination is concerned with the arguments, where conjoint determination is concerned with argument places.)

Redundant determination: an argument x_i partially determines result y of $\langle x_1, \dots, x_i, \dots, x_n, y \rangle \in F$ iff first, the argument place i partially determines y of F , and second, there exists no tuple $\langle x_1', \dots, x_i', \dots, x_n', y' \rangle \in F$ such that their arguments $x_1 - x_n$ and $x_1' - x_n'$ are identical save for x_i and x_i' , and they differ in their result ($y \neq y'$). That is, the argument place plays a role in determining the result y , but it is not the case that changing the argument can change the result.

Russell explains the forms of determination with examples from arithmetic. In the multiplication $0 \times 2 = 0$, the two argument places conjointly determine the result, and each argument place partially determines it. However, the argument value 0 fully determines the result 0, and the argument value 2 only redundantly determines it.

The truth by convention that the deflationist rejects is full determination by the meaning of the sentence. The sentence “all bachelors are unmarried” is made true both by the meaning of the sentence and by the way the world is. It is true because it describes all of the bachelors as being unmarried, and indeed, all of the bachelors are unmarried. However, just as when one multiplies a number by 0 the result is always 0 (despite the fact that the other number is still needed for the result), the meaning of the sentence (allegedly) fully determines the truth of the sentence, and the marital status of the bachelors only redundantly determines it. No matter who the bachelors were, this sentence would still come out as true. Even when considering any possible combination of people who might have been the bachelors, the sentence will still be true, because the meaning of the term ‘bachelor’ dictates that for anything to satisfy it, that thing must be unmarried. So whilst the marital status of the bachelors does play a role in determining the truth of the sentence, it is the linguistic conventions governing meaning that are doing the important work, just as 0 is doing the important work in any multiplication it features in.

The kind of truth by convention that is occurring in deflationism is not of this kind. That $\text{nec}(P)$ is true is not fully determined by the conventions that govern modality, those conventions, along with what is the case at various worlds, conjointly determines the truth of $\text{nec}(P)$. As such, the notion of truth by convention that is employed by the neo-conventionalist is not the same as the problematic kind employed by the positivist. Furthermore, it seems to be an innocuous form of determination that the deflationist should not be troubled by. So, whilst the deflationist position may be described as relying on truth by convention, it is not truth by convention in any problematic sense.

Having presented the deflationary and Humean positions, and examined the role convention plays in those theories, I now turn to a significant criticism of

conventionalism. I present the contingency problem and address the problem as it applies to neo-conventionalist positions.

4. The Contingency Problem

4.1 Setting up the Problem

Any reductive conventional account of modality must face the contingency problem. This is the claim that the contingency of our linguistic conventions jeopardises the necessity of the necessities they determine.⁵⁷ There are several ways of characterising the problem, but the primary tension stems from the desire to maintain:

(CONTINGENCY) We could have held conventions other than those we actually do.

And the validity of the characteristic axiom of the S4 system:

(S4) if $\text{nec}(P)$ then $\text{nec}(\text{nec}(P))$.⁵⁸

It is debatable whether the conventionalist should be required to maintain the S4 principle. Indeed, both Wright and Hale suggest that to insist upon S4 is to beg the question against the conventionalist.⁵⁹ Lewy argues that the contingency of $\text{nec}(P)$, jeopardises the necessity of P .⁶⁰ However, this train of thought requires it to be the case that if possibly possibly P , then possibly P (because if contingently necessarily P then possibly not necessarily P , and if possibly not necessarily P , then possibly contingently P , and so possibly possibly not P), which is a mere transformation of S4, and as such, begs the question against the conventionalist.⁶¹ Using possible worlds semantics, this line of argument assumes the transitivity of the accessibility relation, which is not a safe assumption in the present context. We lack a compelling

⁵⁷ Here I intend 'determine' to be a weak/neutral term for whatever role our conventions play in modality.

⁵⁸ That is, that the S4 axiom is true in every model (where a model is a complete description of the pluriverse).

⁵⁹ Hale (2002), Wright (1980,1985)

⁶⁰ Lewy (1976). Van Cleve (1996) and Sidelle (2009) also argue for this. See also Elder (2006), for a related criticism of conventionalism.

⁶¹ Wright (1985), p.186

argument to force the conventionalist to accept S4. However, theoretical considerations motivate its acceptance. For instance, S4 simplifies our modal logic, and limits the number of distinct possibilities. In S4, possibly possibly P and possibly P are necessarily equivalent, and thus do not represent logically distinct possibilities. With this in mind, for the purposes of this discussion I will assume that if the conventionalist can maintain the validity of S4, then they should.⁶² I consider acceptance of CONTINGENCY to be a safe assumption, as rejection of it would mark too great a deviation from our standard modal beliefs.

The goal of conventionalism is to present a sensible antirealist theory of modality. Antirealist in that it maintains that modality is ultimately explained in terms of our conventions, and sensible in that (among other things) it does not commit one to the truth of the kinds of counterfactuals that require a model that acts as a countermodel to the validity of S4. This can be done by showing that there is a model of conventionalism where S4 is true. Adopting such a model would mean that the conventionalist can maintain the validity of S4.

In this section I present the contingency problem as relevant to both the traditional conventionalist, and the neo-conventionalist. I examine a response from Einheuser that builds upon a response from Wright. I show that the Einheuser response does more to accommodate the conventionalist's modal intuitions, but that it does not fully satisfy some further conditions that ought to be laid upon a response. I then suggest how the response can be revised so as to satisfy these conditions. The resulting model of conventionalism is compatible with the validity of S4, and suitably in the spirit of conventionalism.

Cameron characterises the contingency problem as stemming from the truth by convention of propositions such as <necessarily, P>, and <there are necessary truths>.⁶³ As such, he claims that the neo-conventionalist need not reject S4, because

⁶² For the reader who finds this unacceptable and insists that the conventionalist ought to reject S4, consider this an exercise in mapping the logical space of possible positions.

⁶³ Cameron (2010), p.11 This is in fact a mild mischaracterisation. Cameron gives this reasoning as motivation for the conventionalist to sit on the contingency horn of Blackburn's dilemma, according to which the source of necessary truth must itself be either necessary or contingent. If the source of necessity is itself necessary, then regress ensues; if the source of necessity is contingent, then the

they do not rely on truth by convention (of either the claims necessitated, or the necessitation claims). However, whilst the truth of these propositions by convention is sufficient for the problem to take hold, it is not necessary.⁶⁴ All that is necessary is that the necessitation claims be in some way determined by convention. For this to be the case it is sufficient that (something like) the following counterfactual be true:

(A) If our conventions had been suitably different, then a different range of necessitation claims would have been true.

We can show that A is inconsistent with S4 with the following argument:

(VS4) The axiom scheme $\text{nec}(P) \rightarrow \text{nec}(\text{nec}(P))$ is valid.

(VS4*) Any proposition that is necessary is necessarily necessary. (Informal gloss on VS4)

(A) If our conventions had been suitably different, then a different range of necessitation claims would have been true.

(CONTINGENCY) We could have held conventions other than those we actually do.

(5) If a different range of necessitation claims were true, then there would be at least one necessitation claim that is actually true that would be false.⁶⁵

(6) Let Q be an arbitrary claim that is necessary but would not have been necessary had our conventions been different. (That there should be such a claim follows from 5)

(7) Q is necessary but it could have been that Q was not necessary. (From 6, CONTINGENCY)

necessity of those necessitation claims “has not been explained or identified, so much as undermined” (Blackburn (1993), p.53). However, from his arguments I take it that he thinks these considerations would be sufficient to dispel the contingency problem (so characterised).

⁶⁴ There are some formulations of the contingency problem for which truth by convention is a necessary condition. Van Cleve (1994) gives a particularly strong formulation of the problem, and a further argument for the conclusion that everything is contingent. However, Van Cleve’s argument relies on truth by convention in order to apply to conventionalism. The neo-conventionalist is not threatened by the argument because they reject truth by convention. I will not consider Van Cleve’s argument further here because the focus of this discussion is the threat to neo-conventionalism.

⁶⁵ It might be the case that the difference alluded to in A is that there is a necessitation claim P that is actually false that would be true, but since this would render the actually true necessitation claim $\text{nec}(\neg \text{nec}(P))$ false, I take it that either way the result would be as described in 5.

(8) $\neg(\text{nec}(Q) \rightarrow \text{nec}(\text{nec}(Q)))$ (Formal gloss on 7).

(9) The axiom scheme $\text{nec}(P) \rightarrow \text{nec}(\text{nec}(P))$ is valid and invalid (contradiction from 8 (because 8 provides a countermodel to S4) and VS4 (which insists that there are no countermodels to S4)).

If the conventionalist maintains A then they are committed to a countermodel to S4, and must reject its validity. As such, there is reason for the conventionalist to reject A. However, there is a sense in which the conventionalist wants A to be true, because without the truth of something akin to A, it is not clear how convention could be playing a role in determining necessity (whatever role that might be). This is in violation of LINK:

(LINK) There must be a clear link between our conventions and the true necessitation claims.⁶⁶

Maintaining (something like) LINK should be considered essential to any conventionalist position. A conventionalist account that violates LINK (or an equivalent principle for conventionalisms about realms of discourse other than the modal) cannot be said to be a conventionalism at all, for the conventions cannot be said to be doing the theoretical work in determining the range of modal facts. In light of the desire to maintain S4, and the importance of LINK, the contingency problem can be characterised as the desire for both A and $\neg A$ to be true. We want A to be true because it gives us a link between our conventions and the true necessitations claims (satisfying LINK), and we want $\neg A$ to be true because A commits the conventionalist to the invalidity of S4, as it provides a model where some claims necessitated are only contingently necessary. Wright responds to this problem by suggesting that our modalising is governed by a convention such as:

(RIGIDITY) What it is true to say *of* a hypothetical state of affairs, and what it is true to say *in* a hypothetical state of affairs, is to be determined by reference

⁶⁶ Exactly how LINK should be phrased may be debatable, but something like it should hold. Einheuser insists that conventional features of reality must systematically covary with changes in our conventions. Sidelle (2009) p.231 also appears to recognize this need. LINK is in a similar spirit to this.

to our actual linguistic conventions, even if those are not the conventions that would then obtain.⁶⁷

Wright motivates RIGIDITY with several examples that he takes to be uncontroversial, such as “[i]f ‘red’ had meant what ‘blue’ now means, the sky would have been blue”.⁶⁸ RIGIDITY allows the conventionalist to maintain S4, as even when we consider worlds where the inhabitants hold different conventions, we still make modal judgments relative to our own conventions. However, this is not the whole story. If we take RIGIDITY as the entire response then we still fail to respect LINK. RIGIDITY seems to leave no room for covariance across worlds between conventions and the range of true necessitation claims. This means that there is no sense in which A is true, because no matter what hypothetical situation we entertain, even one in which we have different conventions, the same range of necessitation claims will be true, rendering A false. It was the truth of A that was allowing us to maintain LINK (because if A is true then there is at least some covariance between our conventions and which necessitation claims are true, providing a link between them), so maintaining RIGIDITY is in conflict with LINK because if RIGIDITY is true then A cannot be, and if A is not true then that removes the only support we had for LINK. However, it is important that the conventionalist maintain LINK, as without it there is no reason to think that our conventions actually play a role in determining the necessitation claims. As such, the conventionalist must take further steps to either allow A to be true in some sense, or to secure the truth of LINK by some other means.

Wright might claim that RIGIDITY does not actively conflict with a claim of covariance between our conventions and the necessitation claims, but rather *masks* that covariance.⁶⁹ As such RIGIDITY does not compel us to reject LINK. However, RIGIDITY does seem to conflict with A, whilst it may not give the lie to LINK, further support for it still must be found. This is the task that I take Einheuser to embark upon.

⁶⁷ Wright (1985), p.190

⁶⁸ Wright (1985), p.190. For further discussion motivating RIGIDITY, see Sidelle (2009).

⁶⁹ He says something to this effect in (1985), p.192

4.2 The Einheuser Framework⁷⁰

Einheuser introduces a framework by which we can provide a response in the spirit of RIGIDITY, whilst respecting LINK. She distinguishes between our *conventional practices* (the behavioral regularities that constitute our conventions) and the *carvings* (the theoretical counterpart to our conventional practices) that correspond to them.⁷¹ Based on this distinction Einheuser then introduces a possible worlds framework by which we can model conventional possibility.

There are two kinds of feature a world might have. S-features are those features of a world that are determined by the ‘pre-conventional’ world (i.e. features that are in no way determined by convention). C-features are those features of the world that result from the application of conventions to the world. In the case of conventionalism about modality, the c-features will be (for the neo-conventionalist) those features that correspond to the necessitation claims (or, for the traditional conventionalist, the claims necessitated). The s-features will be all non-modal features of the world (assuming for simplicity that one is conventionalist only about modality).

There are possible and impossible worlds. According to Einheuser, worlds in the framework are ordered pairs of substrata and carvings $\langle s, c \rangle$. A substratum is the sum of the s-features in a world (one should ignore the philosophical connotations that come with the term ‘substratum’, and just take it to be the way the world is, minus those facts that are in some way determined by convention). A carving acts as a function from a substratum to a set of c-features. Einheuser describes a carving as imposing structure on the substratum.

⁷⁰ The semantics Einheuser presents are controversial. However, Einheuser takes it to be justified in responding to the contingency problem. As such, in exploring her response I will assume their acceptability.

⁷¹ For the sake of simplicity, I follow Einheuser in assuming that each world contains only one set of conventional practices. Whilst this is likely false, it will make the following discussion easier, and a full discussion of the consequences of multiple sets of conventional practices is beyond the scope of this discussion.

A world w *supports* a feature F iff there is a carving c such that $\langle s, c \rangle$ contains F (that is, if the substratum contains all s -features necessary for w to have F). A world w *carves* a feature F iff there is a substratum s such that c yields F when applied to s (that is, if the carving in w would yield F in any world that supports F). As such there are two ways that a world could lack some c -feature. The substratum could fail to support it, or the carving could fail to carve it. To use Einheuser's example of ontological conventionalism (concerning mountains), a world could fail to support mountains if the world is flat (so to speak), and has no areas of sufficient elevation, or it could fail to carve mountains if there is no concept of mountain to be applied to areas of sufficient elevation. A world can fail to support a feature even if it carves it, and likewise, can fail to carve a feature even if it supports it.

A world is said to *ground* a carving c when the substratum supports the conventional practices (which are themselves part of the substratum) that correspond with c . There are two ways for the conventions at a world w to differ from those at the actual world $w\alpha$. Either the substratum can fail to support the actual conventional practices, or the carving in w can differ from the actual carving $c\alpha$. In other words, w can fail to ground c , or it can fail to contain c .

There is a world for every substratum/carving combination. As such, the vast majority of worlds contain a carving that does not correspond to the conventional practices of the inhabitants of that world (that is, if there even are any inhabitants). A carving c that is grounded by a substratum s is represented as c_s . Worlds where the substratum of the world grounds the carving of that world are called 'diagonal worlds'.

Since there are two dimensions along which modal change can occur, change in substratum, and change in carving, Einheuser introduces two corresponding kinds of modality: S-modality, and C-modality. S-modality is sensitive only to changes in substratum, whereas C-modality is sensitive only to changes in carving. The domain of S-possible worlds (the S-domain) is the set containing all and only those worlds that contain the actual carving. The domain of C-possible worlds (the C-domain) is the set containing all and only those worlds that contain the actual substratum. The domain that contains all and only the diagonal worlds (the D-domain) overlaps with

both domains at the actual world, and further overlaps with the S-domain at *i*) worlds where the conventional practices remain the same, but other aspects of the substratum differ, and *ii*) worlds where the non-actual conventional practices still ground the actual carving (if there are such worlds). The D-domain further overlaps with the C-domain iff there are worlds where the conventional practices ground more than one carving (which, given our assumption that each world supports only one set of conventional practices, is disregarded), and the actual conventional practices are like this.

In accordance with the spirit of RIGIDITY, the standard modality we use is taken to be a special case of S-modality (the domain of the standard possible worlds is the subset of the S-possible worlds that have substrata that are possible according to the actual world). This means that not only will the vast majority of worlds be non-diagonal, but the vast number of possible worlds as well. RIGIDITY ensures that only worlds with the same carving as the actual world are possible in the standard modality. Since only the actual world and any worlds that also ground the actual carving will be diagonal, and we think it is possible that we could have held different conventional practices, most possible worlds are non-diagonal. Call worlds that are outside the domain of standard modality *præterpossible worlds*.⁷² Any conclusions reached using præterpossible worlds will not be suitable for satisfying standard modal intuitions (as from the perspective of standard modality, præterpossible worlds are impossible worlds). The corresponding notions of necessity for S-modality and C-modality can be defined thus:

NecS(P) at $w = \langle s, c \rangle$ iff for every substratum s' , P is true at $w' = \langle s', c \rangle$ (that is, iff P is the case in every world that differs from w only in its substratum).

NecC(P) at $w = \langle s, c \rangle$ iff for every carving c' , P is true at $w' = \langle s, c' \rangle$ (that is, iff P is the case in every world that differs from w only in its carving).

With these two modal notions available, it becomes clear that conditionals like A are ambiguous. Should they be interpreted as stipulating change along the carving or

⁷² I use this term instead of calling such worlds 'impossible worlds' because they are still possible relative to different modal notions.

substratum dimension? The various disambiguations of conditionals like A can be expressed by the following three types of conditional:

Counterconventional conditional: $P \rightarrow_{cc} Q$ is true at world $w = \langle s, c \rangle$ iff Q at every $w' = \langle s, c' \rangle$ where c' differs from c just enough for P to be true.

Countersubstratum conditional: $P \rightarrow_{cs} Q$ is true at a world $w = \langle s, c \rangle$ iff Q at every $w' = \langle s', c \rangle$ where s' differs from s just enough for P to be true.

Diagonal conditional: $P \rightarrow_d Q$ is true at $w = \langle s, c \rangle$ iff Q at every world $w' = \langle s', c' \rangle$ where s' differs from s just enough for P to be true.

Depending on whether we interpret the term ‘conventions’ in A as meaning conventional practices, or carving, we can interpret A as either a countersubstratum, or counterconventional conditional. As a countersubstratum conditional, A is false. No world $w = \langle s', c \rangle$ where only our conventional practices differ is a world where a different range of necessitation claims are true. This is because there are only two forms of change that can affect what necessitation claims are true. The first is change in carving (which the conditional stipulates against). The second is change in the claims necessitated (for instance, if it were no longer true that Hesperus is identical to Phosphorus), which again is stipulated against in the conditional, which singles out worlds the substrata of which have changed just enough for the conventional practices to differ (and this would represent additional differences beyond mere difference in conventional practices). As a counterconventional/diagonal conditional, A is true. In all worlds $w = \langle s, c' \rangle$ where $c' \neq c$, a different range of claims necessitated are true.^{73, 74}

The falsity of countersubstratum A means that it poses no threat to S4, as it cannot play the role of A in deriving the contradiction 9 in the argument above. Counterconventional A is true, but is of no use in the derivation of 9 because the worlds that make counterconventional A true are C-possible but not S-possible (and

⁷³ Because carvings are functions, any two carvings that produce the same range of true necessitation claims (from the same s -features) are identical.

⁷⁴ Diagonal A is also true because all diagonal worlds with different conventional practices will be counterconventional (on the assumption that different sets of conventional practices cannot correspond to the same carving).

since our standard modality is a special case of S-modality, those worlds are präterpossible, and so not relevant to the validity of S4, which is an axiom of modal logic applied to standard modality). This means that the possibility of a different range of necessitation claims being true is mere C-possibility, and not the possibility of standard modality. As such, counterconventional A poses no threat to S4, which is a principle of standard modality.

Understood in this way, the conventionalist is not committed to any countermodels to S4, because the problematic countersubstratum A is false, and the worlds required for the truth of counterconventional A do not feature in the models relevant to the validity of S4, as they are präterpossible to standard modality. However, the truth of counterconventional A provides support for LINK because it allows there to be a sense in which A is true, showing that there is a clear link between our conventions (interpreted as carvings) and the true necessitation claims. Further, that our standard notion of modality falls within S-modality shows us why we are inclined to think that RIGIDITY is true (though it is worth noting that RIGIDITY is seemingly only de facto true for standard modalising, not de jure, because it just happens that no possible worlds contain non-actual carvings).

The Einheuser framework allows us to present conventionalism in a way that bears no commitment to countermodels to S4. This allows us to maintain the kind of sensible antirealism mentioned earlier.

4.3. Applying the Einheuser Framework to Neo-Conventionalism

The Einheuser framework is compatible with neo-conventionalist accounts of modality. The notion of a substratum in a world is compatible with various possible world accounts. It could be a set of sentences, a state of affairs, or something physical, more akin to a Lewisian world. For the neo-conventionalist, it is important that all of the claims necessitated that are true at a world w are made true by the substratum in w .

More difficult is determining how one interprets the carving as taking one from a substratum to the range of necessitation claims. This cannot be done directly. If it were, then it would be the case that, whilst the claims necessitated are not true by convention, the necessitation claims are. This is a claim that the neo-conventionalist is eager to avoid.

To accommodate the Humean, the function c takes one from the substratum to a list of different 'kinds of claims', e.g. mathematical, metaphysical, etc. Any true propositions P of these kinds will then have a corresponding true necessitation claim $\text{nec}(P)$. $\text{Nec}(P)$ is not true by convention, but true because P is true and of a certain kind (this at least is the Humean line, so if it is acceptable there, then it is acceptable here).

To accommodate the deflationist, the function c takes one from a substratum s to a set of worlds that are accessible from the world containing s . There will be a true necessitation claim $\text{nec}(P)$ for every proposition P that is true at every world that is accessible according to c . $\text{Nec}(P)$ is not true by convention, but because P is true at every possible world.

In doing this we are using conventions to select between different candidate notions of necessity, rather than making any individual necessitation claims true by convention. The claims necessitated we conventionally select are already true (by virtue of being true at every 'possible world', for some meaning of possible world), regardless of our conventions, what the conventions are doing is selecting one candidate notion of necessity to be of interest. In doing this we select a range of necessitation claims to give special significance to. However, this does not affect the threat of the contingency problem, and does not affect the Einheuser framework's potential as a response (despite Einheuser seemingly not having neo-conventionalism in mind).

4.4 The Contingency Problem: Continued

The Einheuser framework allows us to maintain S4, CONTINGENCY, LINK, and A (and on a different reading, $\neg A$), by distinguishing between our conventional practices and the carvings that correspond to them. By identifying an ambiguity in A, and exploiting the possible disambiguations, Einheuser is able to satisfy the *prima facie* conflicting intuitions. This allows us to present a model in which conventionalism is true, but that is not contrary to the validity of S4. However, in the Einheuser framework LINK is true relative to counterconventional A, but false relative to countersubstratum A. This is problematic because it is not enough that our carving should determine the range of true necessitation claims, it must also be the case that our conventional practices themselves determine our carving (and thus, indirectly, the necessitation claims). Einheuser maintains a covariance claim for carvings at the expense of maintaining one for conventional practices, but any conventionalist must be able to identify that our conventional practices (and not just the carving that corresponds to them) play a role in determining *c*-features. In other words, the conventionalist should want to maintain:

(LINK') There must be a clear link between our *conventional practices* and the true necessitation claims.

As it stands, it is the carving that is doing all of the theoretical work. This, combined with the fact that the vast majority of worlds are non-diagonal, suggests that the conventional practices held at a world play no role in determining the carving in that world. This violates LINK'. LINK' could be maintained if a conditional like B were true.

(B) If a world with suitably different conventional practices *were actual*, then a different range of necessitation claims would have been true.⁷⁵

B would preserve LINK' because even though carvings do not correspond to conventional practices in most worlds, if any world *w* were actual, then there is

⁷⁵ By 'actually' here I do not mean the rigidifying actuality operator signified by @, but rather actuality as a status that a world may or may not have.

reason to claim that the conventional practices and the carving in that world would match. We want it to be the case that the conventional practices in the actual world determine the actual world's carving. Otherwise it is not clear that the carving has anything to do with convention, and that it is not just some philosophical trick brought into the theory to avoid a problem. B is different from A because it specifies conventional practices, not just conventions, and because the antecedent stipulates that such a world be actual. When we consider a world where we hold different conventional practices (as A invites us to do under one reading), the same necessitation claims hold (because of RIGIDITY). When we consider if such a world had been *actual* (as B invites us to do), different necessitation claims hold, because RIGIDITY stipulates that we consider hypothetical situations relative to our actual conventions.

If some rule like B were not in effect, then we would be left with the awkward conclusion that it is mere coincidence that the actual world is diagonal, and the worrying epistemic question of how we even know that we live in a diagonal world. As such, at least in the actual world, the carving should be determined by the conventional practices, establishing a clear link between the two; this is what B would do.

How does B translate into the three forms of conditional introduced above? Where w' is a world that supports different conventional practices and Φ_w is the set of necessitation claims true at w :

Countersubstratum B $\langle w' \text{ is actual} \rightarrow_{cs} \Phi_{w'} \neq \Phi_\alpha \rangle$ is true at world $w = \langle s_\alpha, c_\alpha \rangle$ iff $\Phi_{w'} \neq \Phi_\alpha$ at every $w'' = \langle s', c \rangle$ where s' differs from s_α just enough for w'' (to have different conventional practices and) to be actual.

Counterconventional B $\langle w' \text{ is actual} \rightarrow_{cc} \Phi_{w'} \neq \Phi_\alpha \rangle$ is true at world $w = \langle s_\alpha, c_\alpha \rangle$ iff $\Phi_{w'} \neq \Phi_\alpha$ at every $w'' = \langle s, c' \rangle$ where c' differs from c_α just enough for w'' (to have different conventional practices and) to be actual.

Diagonal B $\langle w' \text{ is actual} \rightarrow_d \Phi_{w'} \neq \Phi_\alpha \rangle$ is true at world $w = \langle s_\alpha, c_\alpha \rangle$ iff $\Phi_{w'} \neq \Phi_\alpha$ at every world $w'' = \langle s', c_s' \rangle$ where s' differs from s_α just enough for w'' (to have different conventional practices and) to be actual.

Countersubstratum B claims that B is true iff the set of true necessitation claims is different at every world w where the substratum at w differs just enough for it to be the case that w is the actual world. Counterconventional B claims that B is true iff the set of true necessitation claims is different at every world w where the carving at w differs just enough for it to be the case that w is the actual world. Diagonal B claims that B is true iff the set of true necessitation claims is different at every diagonal world w where the substratum at w differs just enough for it to be the case that w is the actual world.

There are two barriers to any of these conditionals being fit to play the role B was introduced for. First, all require that there should be some change in either the carving or substratum of a world that would make that world actual. It is not clear in any possible worlds account what change could take place inside a world that would make it the actual world. Second, even if this can be overcome, each conditional has its own flaws that make it unusable.

First consider countersubstratum B. The antecedent of B stipulates substratum change, change in conventional practices. For the truth of the antecedent to be able to make the consequent true, the world in question must be diagonal, and since the only S-possible diagonal worlds are the actual world and those with identical conventional practices (and thus won't make the consequent true),⁷⁶ the antecedent of B is S-impossible, and as such B is trivially true. This is because the antecedent requires change along one dimension (change in substratum), whilst the consequent requires change along a separate dimension (change in carving). Since the C-domain and the S-domain only overlap at the actual world, and the actual world is expressly excluded from this consideration, countersubstratum B can only be trivially true. Countersubstratum conditionals, as a part of S-modality, can only function along one dimension of change, and thus are of no use to us here. Likewise, since the antecedent of B stipulates difference in conventional practices (part of the substratum) and not in carving, there can be no C-possible worlds that satisfy the

⁷⁶ Assuming that only one set of conventional practices can ground each carving. Whilst this is probably false, it simplifies the discussion and its falsity would not affect the result.

antecedent, rendering counterconventional B trivially true as well. Diagonal B would come out true, but the worlds that satisfy the antecedent are präterpossible worlds and as such, do not satisfy the intuitive motivation for B. Even before distinguishing between the different conditionals, the conventionalist should be inclined to accept B as a standard counterfactual conditional (which Einheuser showed us is a special case of the countersubstratum conditional). The tension that made the conventionalist want to reject A does not apply to B, because B does not jeopardise the validity of S4. B does not threaten S4 for the same reason that counterconventional A does not threaten S4. S4 is a principle of standard modality, the modality based on the carving in the actual world. The antecedent of B identifies a non-standard modality, one based on the carving of a non-actual world. As such, the worlds that make B true are präterpossible relative to S4, and so cannot pose it a threat (because nothing that is the case at worlds outside of those in the domain of standard modality can create a countermodel to the principles of standard modality). That is, for any proposition (true as a result of B) that says that the claim necessitated P might not have been necessary, the possibility in that proposition is not the possibility of standard modality. So, if one of the Einheuser conditionals is going to be true then it needs to be countersubstratum B. How can B be (non-trivially) true?

4.5 An Unsuccessful Response

One ultimately unsuccessful attempt at a solution might be to introduce a rule like:

(DIAGONAL) The actual carving must correspond to the conventional practices in the actual world.

However, this is equivalent to:

(DIAGONAL') The actual world must be a diagonal world.

violating the principle: ‘if possibly P then possibly actually P’.⁷⁷ To insist that the actual world must be a diagonal world would be to insist that no non-diagonal world is a possible world, since if to be possible is to be possibly actual, then it must be the case that any possible world could have had the status of actuality, diagonal or not. To insist that whichever world is actual must be diagonal is to deny that there are non-diagonal possible worlds.

4.6 The Proposed Solution

As it stands, the Einheuser framework is unable to maintain LINK’. Here I suggest that the conventionalist can maintain the utility of the Einheuser framework, whilst respecting LINK’, if they abandon the notion of worlds being ordered pairs of substrata and carvings. Instead, the conventionalist can identify worlds with substrata alone, and when modalising, consider propositions to be true or false at a world relative to a carving. Einheuser assents to the viability of such a position. She claims that nothing of substance hangs on such a change, because “[a]nalogues of all of the concepts [she] introduce[s ...] can be introduced for this construal of conventionalist possible worlds and the accompanying relative notion of truth at a world.”⁷⁸ First, one must present analogues for the three Einheuser conditionals:

Counterconventional conditional: $P \rightarrow_{cc} Q$ is true at world w relative to c iff Q at w relative to c' , where c' differs from c just enough for P to be true relative to c' .

Countersubstratum conditional: $P \rightarrow_{cs} Q$ is true at a world w relative to c iff Q at every w' relative to c , where w' differs from w just enough for P to be true relative to c .

Diagonal conditional: $P \rightarrow_d Q$ is true at w relative to c iff Q at every world w' relative to c' that is grounded by w' , where w' differs from w just enough for P to be true relative to a carving c' .

⁷⁷ Whilst I do not know of anywhere that explicitly endorses such a principle, Wright does suggest as much in (1985), p.186.

⁷⁸ Einheuser (2006), p.463 fn.7

We can also introduce C, a version of B more suitable to the revised Einheuser framework.

(C) If a substratum supporting a suitably different set of conventional practices were actual, then a different range of necessitation claims would have been true.

Whilst the new conditionals are still of use in representing conventional modality, they are of little use in representing C. In trying to secure the truth of C we can make use of the distinction between two suppositional acts: A-supposition (roughly, supposing as actual) and C-supposition (roughly, supposing at counterfactual).⁷⁹ In A-supposing that P, we are acting as if we believe that P is actually the case. In doing so, we ‘jump perspective’ to that of the world(s) under consideration. A-suppositions have long been used to think about conditionals with impossible antecedents (e.g. if water had had the molecular structure XYZ, then it would have been so necessarily). The suppositional acts associated with all of the Einheuser conditionals are C-suppositions, whereas (due to its use of the term ‘actual’) C is clearly inviting one to engage in A-supposition. With this in mind, we can take C to be an A-suppositional equivalent of the countersubstratum conditional interpretation of A. That is, as:

(C’) Under the supposition that we hold suitably different conventional practices than those in the actual world (where ‘the actual world’ rigidly designates the world that is actual outside of this supposition), a different range of necessitation claims are true.

One can then maintain a principle like:

(**RIGIDITY**) The carving relative to which we judge worlds must correspond to the conventional practices held in the actual world.

RIGIDITY’ ensures that C’ is true, and C’ is sufficient for the conventionalist to maintain LINK’.

⁷⁹ The distinction is further explained in Chalmers (2002), Divers & Elstein (2012), Divers & González-Varela (2013).

At this point the reader might wonder why such a move was not made earlier, to rescue the Einheuser framework before making the above changes to it. This may seem like a promising move. However, in order for an A-supposition of B to be true, it would still be required that we hold some principle akin to DIAGONAL. As we saw earlier, a principle along these lines will lead to radical departure from modal orthodoxy. In order for it to be the case that the actuality of a world with different conventional practices would result in the carving corresponding to those practices also being actual, it would need to be the case that every such world is already diagonal. However, when considering worlds as substrata, and considering truth at worlds relative to a carving, no such constraint need be placed on the worlds under consideration. This eliminates any threat to the possible actuality (and thus the possibility) of any possible worlds. A-supposition of C requires only a principle like RIGIDITY'. RIGIDITY' is similar enough to RIGIDITY (it is little more than phrasing RIGIDITY in the terms of the revised Einheuser framework) that it can be motivated by the same considerations that enabled Wright to present it.

The reader might also wonder as to the point of this endeavour if the final result so closely resembles Wright's original position. There is however a key difference, in that this position allows the conventionalist a response to the original conflicting intuitions about claims such as A. Under the current position, the conventionalist can maintain that A is both true and not true (under counterconventional and countersubstratum interpretations respectively). It may seem counterproductive to make such changes to a position for the mere preservation of intuitions, but as was shown earlier, failure to respect either part of this intuition results in failure to maintain LINK (and in turn, LINK'), which is a serious problem for a conventionalist.

4.7 Guarding Against the Original Contingency Threat

There is a risk in this approach that by maintaining RIGIDITY' and C, the position is re-exposed to the threat originally posed by A. We must ensure that the truth of C

does not jeopardise the validity of S4. Previously we had the ambiguity between the countersubstratum and counterconventional readings of the conditional to guard against this, but with no (obvious) viable counterconventional reading of C, this option is not available. However, we can use the tools of A-supposition and C-supposition to do the same work with C that Einheuser uses counterconventional and countersubstratum conditionals for. The C-supposition of C is not true, and as such does not commit the conventionalist to a countermodel of S4. The A-supposition of C (that is, C') is true, but provides no countermodel to S4 because the antecedent of the conditional is präterpossible (in that, when combined with RIGIDITY', it invites us to 'jump perspective', thus placing us outside of the domain of standard modality). A-supposition of C poses no threat to S4 in much the same way that counterconventional A did not pose a threat to S4. It is made true by worlds relative to a präterpossible carving that is not relevant to axioms of modal logic for standard modality, such as S4.

Conventionalism, as presented here, bears no commitment to countermodels to S4. Of the conditionals considered here, those that are problematic (that would require that the model be a countermodel to S4) are false, and those that the conventionalist needs to maintain are not problematic (they are made true by worlds that do not render S4 false in the model, and so do not make it a countermodel). This means that there are options available to the conventionalist that wants to maintain the validity of S4, leaving the ground level for arguments as to whether or not maintaining S4 is something the conventionalist ought to do. That previous responses to the contingency problem have resulted in failure to fully respect the role of convention in an account of modality is symptomatic of a common neglect to detail as to how the conventions play the role that they do. Whilst it may not be too problematic in some instances to lose track of exactly *how* our conventional practices work in determining modality, it is a critical problem when it allows the conventionalist to be unable to demonstrate *that* they do. The revised Einheuser framework at least satisfies this weaker condition.

5. Two Unanswered Questions for Deflationism and Humeanism

So far the goal of the chapter has been to specify exactly what role convention plays in the neo-conventionalist account. With this done there are two significant deficits in the account that need to be explored. These concerns stem from the lack of explanation as to why we have conventions that govern modality in the first place, and by what conventional practices our modalising is explained. When we give a conventional explanation for anything, we need to be able to pinpoint which of our conventional practices do the work, elucidate how those conventions explain the desired phenomenon, and state why people hold those conventions. For instance, one might give a conventional explanation of why (in the UK) it is illegal to drive on the right hand side of the road. One can explain how the convention makes it the case that driving on the right is illegal, but there must also be an explanation of why we decided to have such conventions. This explanation cannot simply be along the lines of accounting for the truth of certain sentences, or validating a certain kind of discourse. The desired explanation is of why the convention users hold the conventions, not why the philosopher theorises about them. In the driving example this explanation is easy enough. We want to avoid head on collisions on the roads, so by making it illegal to drive on the right, everyone will drive on the left, making such collisions less likely. Any kind of conventionalist position owes such an account.

This demand is equivalent to one that Cameron makes of those who think that there is a metaphysical distinction between the possible and impossible worlds. Cameron uses the analogies between time and modality to make this demand as part of his case for deflationism.⁸⁰ In particular he considers the difference between how the A and B-theorists understand the present. The A-theorist thinks that there is a natural distinction between the present and non-present. The B-theorist does not. Cameron draws a parallel between this and the distinction between the possible and impossible worlds, the deflationist being analogous to the B-theorist. Cameron claims that the A-theorist is bound to provide an explanation of what it is that makes the present time privileged. He demands, and rightly so, that the A-theorist be able

⁸⁰ Cameron (2010a), p.10

to identify a metaphysical reason that distinguishes the present from other times. He is right to carry this commitment across the analogical barrier to the modal case. Those who claim that there is a natural distinction between the possible and impossible worlds owe us an explanation of what metaphysical difference determines that distinction. However, Cameron is not right to claim that the deflationist does not owe us an equivalent explanation. Of course, this explanation is not going to be metaphysical, to ask for a metaphysical explanation from the deflationist would be question begging. The explanation owed is a conventional story that indicates why we put the distinction where we do (and indeed, why we have a distinction at all), and by what conventional practices we place it there. This may be an explanation in terms of conceivability, instrumental utility, or perhaps something else. What is important is that such an explanation is explicitly made. The deflationist cannot simply wave their hand and say “the distinction is merely conventional”. This would be equivalent to the A-theorist saying “the present is privileged for metaphysical reasons”. We don’t accept this answer from the A-theorist; we should not accept it from the deflationist. Note that the B-theorist, of whom the deflationist is supposed to be an analogue, does face the same challenge in explaining the difference between the present and the non-present. The difference is that the B-theorist provides an explanation. Just as the deflationist’s explanation doesn’t need to be metaphysical, neither does the B-theorist’s. The B-theorist explains that the present has privileged status because it is where we are. Granted, it is a simple explanation, and we usually accept the dismissive claim that the explanation is ‘because of our perspective’ because it is obvious what is meant by this. However, when dismissing the explanation of the distinction between the possible and impossible worlds by saying that the explanation is conventional there is no equivalent transparency, no more than if the A-theorist were to say that the difference is simply metaphysical. When describing the actualist who thinks that the distinction is natural, Cameron says: “It’s as if we’re saying that some of the worlds come with a special glow – the glow of possibility – but we’re not saying anything about what this glow is, or how we can detect it; and that’s just bad

metaphysics.”⁸¹ This is true, but it cuts both ways. If the deflationist doesn’t give an explanation of the ‘conventional glow’ that the possible worlds have, then they are guilty of the same sin.

No such explanation is given on the part of the deflationist. Whilst this shows that the account is incomplete, it does not detract from the original insight that motivates the position. Cameron does at least acknowledge the importance of these questions, though he seems to see addressing them as the next big task to be completed.⁸² I think this is getting the cart before the horse. Without an understanding of why we modalise, and an explanation of what our conventional practices are and how they place the distinction where it is, there is no theory, just a statement of what kind of theory might work. It’s no better than saying “the distinction is natural”, without providing any metaphysical explanation. The question remains; what kind of conventional explanation can be given?

6. Conclusion

In light of the problems presented in Section Five, the deflationary and Humean accounts are inadequate for the task of providing a full account of modality. This is not because of any defeating objection, or internal inconsistency, but because they are not complete accounts. It would be uncharitable however to take this as indication of their lack of worth. Rather, we should consider the value of this work as identifying and promoting the neo-conventionalist move, the retreat from type 1 conventional grounding to type 3 conventional grounding. In showing that the role of convention in accounting for necessity is not as simple as originally thought, these positions clear space for other positions that make the same move, and resolve the questions of Section Five. In the next chapter I examine Sidelle’s neo-conventionalism, which does just that.

⁸¹ Cameron (2008), p.277

⁸² Cameron (2010b)

Chapter Two: Sidellean Conventionalism

1. Introduction

In the previous chapter I established that the deflationary and Humean forms of neo-conventionalism, whilst providing an antirealist account of modality, fail to give an explanation of how or why we modalise (or in particular, why we modalise in the way we do).

In this chapter I present Alan Sidelle's neo-conventionalism (though he does not characterise his position as such) about modality.⁸³ I give a full exposition of the account, and show how it can be used to give an antirealist account of essence. I defend it from Gillian Russell's indexicality problem, and I examine just how we should interpret and respond to Sidelle's apparent reliance on analyticity. I conclude that, despite appearances, the account does not rely on analyticity. I show that Sidellean conventionalism is better placed to account for essence than the deflationary or Humean neo-conventionalisms, in part because it provides a better antirealist account of necessity, and in part because Sidelle's account bridges the gap between the modal and neo-Aristotelian theories of essence. I conclude that Sidellean conventionalism is the best chance to provide an antirealist account of essence as understood by the modal essentialist.

2. Sidellean Conventionalism

Sidelle uses the phrase 'necessary truths' to denote the explanandum of his account. This leaves ambiguity as to whether the account should be phrased in terms of sentences, propositions, facts, or something else. Further confusion results from Sidelle referring to the analyticity of some truths. Analyticity being a property of sentences, this suggests that the truths Sidelle is referring to are sentences, yet this

⁸³ Sidelle (1989, 1998, 2010)

seems incongruous with his motives. Presumably (and charitably) the choice of ‘necessary truths’ is intended to remain neutral to the various linguistic and metaphysical distinctions relevant to necessity (much in the same way that I use ‘claim’ as neutral between various linguistic forms and the metaphysical states they represent). Furthermore, in using ‘necessary truths’ Sidelle is subject to the ambiguity resolved by distinguishing between necessitation claims and claims necessitated in the introduction to this thesis. In the expository sections of this chapter I present Sidelle’s position in terms of necessitation claims and claims necessitated, predicating analyticity to them in the way that Sidelle does to necessary truths.

Sidelle writes in response to the challenge presented to traditional conventionalism by necessity a posteriori. In doing so he allows us to give an antirealist account of modal essentialist predication by which essential facts reduce to modal facts, which in turn are conventional in the way Sidelle sets out.

Traditional conventionalism is typically characterised as the claim that all necessity has its source in our linguistic conventions; that any necessitation claim $\text{nec}(P)$ is true because the corresponding claim necessitated P is analytic. Necessity a posteriori is considered to be problematic for traditional conventionalism because claims necessitated that are analytic can be known a priori, whereas claims necessitated that are a posteriori require (*ex hypothesi*) some kind of empirical discovery, and so cannot be known a priori. The important characteristic of necessities a posteriori is that they are epistemically contingent. That is, given the empirical nature of any a posteriori claim necessitated, for example that the elemental composition of water is H_2O , we can easily imagine that it might not have been the case; that the elemental composition of water might instead have been discovered to be XYZ. As such, it cannot be known a priori, and that which cannot be known a priori cannot be analytic.

Therefore, by the argument:

- (1) Only analytic claims can be claims necessitated. (Traditional conventionalism)

- (2) If a claim is analytic then it must be knowable a priori. (From above reasoning)
- (3) A posteriori claims cannot be known a priori. (Definitive of 'a posteriori')
- (4) Therefore a posteriori claims cannot be analytic. (Modus tollens from 2 and 3)
- (5) Therefore a posteriori claims cannot be claims necessitated. (From 1 and 4)

The conventionalist might conclude that there is no necessity a posteriori. However, to do so would be unacceptably revisionary for this project. Instead, Sidelle rejects premise 1, concluding that the identification of necessity with analyticity is not essential to the conventionalist position.

It is important to note that for a necessitation claim to be a posteriori, it is sufficient that the claim necessitated be a posteriori. The source of the a posteriority of the necessitation claim is only in the a posteriority of the claim necessitated. Having established the truth of the claim necessitated, it is not a matter of further empirical investigation to know that the necessitation claim is true. Indeed, it is not clear what further empirical investigation could take one from the truth of empirical claim P to the necessity of P. Because of this, Sidelle divides a posteriori necessitation claims into two components: one empirical and contingent, one conventional but that bestows the claims necessitated with their modal status. Sidelle claims that it is not a problem for the conventionalist if the a posteriori claim necessitated is not analytic, so long as its modal status comes from convention. This is what he sets about demonstrating.

To Sidelle, "convention [is just] a catch-all for mind-based contribution".⁸⁴ As such, this account need not be understood in terms of actual conventions, or taken along with any theory of convention. As commonly understood, for something to be conventional there must be an element of choice involved, whether this choice is explicit or implicit. Furthermore, it must be the case that there are alternative choices that might have been made that would have been just as good for the convention makers. Sidelle's understanding of convention is much weaker than this. For

⁸⁴ Sidelle (1989), p.2

something to be conventional it need only be determined by the way we think or act in a certain context. This mind-based contribution could be something that we have no choice about.⁸⁵ It could be pre-determined by the way our brains work, or there could be no alternative that would yield practicable results. So long as it is our interest relative interactions with the world that are doing the work, and not the world itself, this still counts as conventional for Sidelle. This means that we need not think of the account as actually being strictly conventional. The framework that Sidelle lays out can be used for a non-conventional antirealism. However, for the sake of simplicity I will continue to talk of the position as a conventionalist one.

3. General Principles of Individuation

Sidelle introduces general principles of individuation (GPIs). GPIs are the form conventions take in the mechanics of Sidelle's account. They can be schematically presented as:

GPI Schema: (x) (If x belongs to kind K , then (if p is x 's P -property, then it is necessary that x is p)).⁸⁶

Where x is a variable for individuals, kinds, or properties; P -property is a kind of property that is taken to be definitive of kind K , and p is a specific instance of that property. For example, if x = water, K = chemical kind, P -property = elemental composition, and p = H_2O , then the resulting GPI would be:

WATER: If water is a chemical kind, then if H_2O is water's elemental composition, then water necessarily has the elemental composition H_2O .

The reasoning that secures the necessity of water having the elemental composition H_2O takes the following form:

⁸⁵ It is not clear to me if Sidelle would accept this particular claim, but I assert it, regardless.

⁸⁶ Sidelle (1989), p.34

WATER: If water is a chemical kind, then if H_2O is water's elemental composition, then water necessarily has the elemental composition H_2O .
(GPI)

P1: Water is a chemical kind. (Premise)

C1: If H_2O is water's elemental composition, then water necessarily has the elemental composition H_2O . (By modus ponens from WATER and P1)

P2: Water has the elemental composition H_2O . (Empirical premise)

C2: Water necessarily has the elemental composition H_2O . (By modus ponens from C1 and P2)

Sidelle's position can be characterised as a weakened form of the traditional conventionalist account. Sidelle retreats from the traditional explanation of necessity (that necessarily P iff P is analytic) to:

A necessitation claim $nec(P)$ is true iff P (the claim necessitated) is analytic, or $nec(P)$ is the conclusion of a valid inference from the claim necessitated and an appropriate true GPI.⁸⁷

GPIs explain how a posteriori necessitation claims are made true, regardless of whether one takes a realist or antirealist interpretation of necessity. To be an antirealist, Sidelle suggests that one interpret GPIs not as facts about the way the world is, but rather as object language formulations of conventions we hold. As such he takes them to be analytic. Just how we should interpret and respond to Sidelle's apparent reliance on analyticity will be addressed in Section Nine. The reader's concerns about the reliance on analyticity, or just what Sidelle means by 'analyticity' should be reserved until then. For now I focus on Sidelle's account.

The GPIs as given above are in fact not the conventions themselves, but the claims that are made analytic by them. The conventions themselves govern how we use our terms for the subjects of the GPIs. Sidelle introduces *kind terms*. Kind terms are terms that we introduce with the intention to use in a way that is governed by the GPIs for a specific kind. A K-term is a kind term of kind K.

⁸⁷ The appropriate GPIs for any claim necessitated are those of the things it refers to. This stipulation ensures that no unwanted consequences are made necessary through logical closure.

Given the way that we apply a kind term, counterfactual applications of that term are governed by certain (conventionally selected) aspects of that original application. The aspects that are selected as important are the P-properties. A kind term L is a K-term if the P-property for the kind K (for instance, elemental composition for a chemical kind) is found in most of the things we apply L to, and governs our counterfactual use of L. In this way P-properties are definitive of kinds because we are resolved not to apply a kind term in situations where the intended referent does not have the P-property for that kind term. For example, ‘water’ is a chemical kind-term because whether we correctly apply the term ‘water’ in any given instance is determined in relation to the P-property selected by the GPI for chemical kinds (in this case elemental composition). The elemental composition of “most (enough) of the items to which we apply (by which we introduced) the term” is H_2O , so we resolve only to apply the term ‘water’ to things which have the elemental composition H_2O .⁸⁸ Likewise, ‘chair’ might be considered a furniture-term, governed by the GPI for furniture; ‘chemical kind’ might be classified as a natural kind-term, and ‘furniture’ might be an unnatural kind-term. We can present the conventional schema for GPIs as:

GPI Conventional Schema: If ‘x’ is a K-term then if p is the P-property of the thing denoted by ‘x’, then ‘x’ applies to something in any possible situation only if it is (has) p .⁸⁹

Our use of a term L is governed by a GPI. The kind of kind term we introduce L as determines which GPIs govern it. Introducing L as a K-term commits us to governing the use of L in terms of the GPI for kind K. This effectively makes L an instance of K, and according to Sidelle this makes it analytic that L is an instance of K. This is the way that, according to Sidelle, we use our conventions to create kinds. By introducing ‘water’ as a chemical kind term we establish that it is analytic that water is a chemical kind.⁹⁰ This means that it is a matter of convention, and not the pre-

⁸⁸ Sidelle (1989), p.45

⁸⁹ Sidelle (1989), p.44

⁹⁰ Sidelle (1989), p.46

conventional world, that water is a chemical kind. 'Water' is introduced as a chemical kind-term, and elemental composition is selected as the P-property for chemical kinds (through our holding the convention that is the GPI for chemical kinds). For this reason, Sidelle claims that kinds themselves are conventional in nature.⁹¹

This is not to say that we are infallible in our kind judgments. There are various ways in which we can be mistaken. It may well be the case that of the first instances of gold that we encountered, the primary exemplars by which we introduced the chemical kind-term 'gold', half of them were in fact fool's gold (FeS_2) and not gold (Au).

The mistake here is not that we recognised there being a kind when there in fact was none, nor is it that we recognised a natural kind when there is only a non-natural one. Whether or not there really is a kind depends entirely on whether gold and fool's gold make a kind together. If there is a kind, then whether or not it is natural depends on whether such a kind fulfils our criteria for being natural. We must take pains to avoid the unsophisticated form of antirealism whereby we talk *as if* there are kinds when in fact there are none.⁹² There are kinds, and we are right to think that there are. The source of the antirealism in the account is that we feature indispensably in the explanation of what it means for us to be right in this regard.

The mistake is that we recognised there being a *chemical kind*, where there could not be one. For there to be a kind *just is* for us to resolve to govern our use of a K-term in accordance with the GPI for K, in conjunction with certain empirical facts (with the GPI determining which facts those are). Introducing 'gold' as a chemical kind-term commits us to counting things as gold or not in accordance with whether those things demonstrate a certain kind of similarity. That is, similarity of elemental composition (the P-property for chemical kinds), whatever that elemental composition may be.⁹³ If our exemplars do not have a common elemental composition then they do not make up a chemical kind, regardless of whether we

⁹¹ This claim is examined more thoroughly in Section Four.

⁹² Blackburn argues against this interpretation of antirealism in Blackburn (1993).

⁹³ For simplicity I ignore the fact that gold was doubtless not originally classified based on its elemental composition.

think they do or not. Just because we are resolved to use a term in a specific way, does not mean that we will always be successful in doing so. What is important is that when we apply 'water' to something, and then discover that that thing is in fact not H_2O , our response is not to conclude that not all water is H_2O , but to revise our judgement and conclude that this is not an instance of water.

Similar rules apply when considering essentialist predication regarding individuals. This requires that we introduce names of individuals as *kind names*. Just as with kind terms, by introducing a name as a K-name, we implicitly determine that talk of the thing in question will be governed by the GPIs for kind K. As such, the origin (for example) of the individual to which the name is applied will be a K-origin. A K-origin is the kind of origin that individuals of kind-K have. This allows us to be certain of the referents of our terms, even when there are multiple overlapping candidates. For example, if 'Socrates' is introduced as a human-name, then we determine that the referent of 'Socrates' is the person, the thing with the kind of origin that humans have (i.e. biological origins) rather than the aggregate of biological matter that happens to make up Socrates at the time of ostention. The former has a human-origin, whereas the latter has only a 'collection-of-matter' origin.

To introduce 'Socrates' as a human-name is to intend to use 'Socrates' in such a way that nothing counts as Socrates unless it corresponds to the subject of our ostention (or the first ostention by which Socrates was originally named) in accordance with our identity conditions for humans. This commits us to not calling anything (in any possible situation) 'Socrates' that does not have, inter alia, the biological origin of the subject of our ostention when we introduced the human-name 'Socrates'.

Once again, this does not require that we be infallible in our judgments concerning individuals. If the thing which we refer to with the term 'Socrates' were, unbeknownst to us, a robot (for the purposes of disambiguation I will call it 'Socratron'), then it would not be the case that by giving it a human-name we bestow it with a biological origin. Human identity conditions fail to tell us if something in another world is Socratron or not. We cannot ask if something has the same mother

as Socratron, for Socratron does not have a mother. The mistake here is a category mistake. When we give something a kind name, we determine which of the features it already has are necessary to it. This is done by placing it within a kind and so determining that use of that name will be governed by the GPIs for that kind. The thing's P-properties are determined by the P-properties of the kind it falls under. If by applying the human-name Socrates to Socratron we commit ourselves to applying that term only in accordance with the identity conditions for humans (which are determined by the GPIs for humans) then we have simply applied a term that by the rules of our language cannot be correctly applied in this way. If we do this in ignorance then this will lead us to have false beliefs about Socratron, but this is exactly what we should expect if it turned out that what we had been calling Socrates was a robot all along.

4. The Link between Modality and Ontology

According to Sidelle, we use the conventions described by the GPIs to carve the world into objects and kinds. Unlike the metaphorical butcher of metaphysical realism, for whom carving at the joints of reality is a matter of identifying metaphysical joints at which to cut, conventional carving is more like that done by the cheesemonger. When we carve the world with our conventions we do so at no metaphysical joints in the world-cheese. There are joints in the cheese, but they are the result of the way we interact with it in accordance with our own interest-relative purposes. The Sidellean conventions that give us modality do so by determining which properties are important to, and definitive of, objects and kinds. This provides us with an explanation of why we modalise in the first place. We need to be able to categorise the world in such a way that we can think about and interact with it in a way that suits us. Through establishing this kind of taxonomy we create the modal facts. Modality is the result of the process by which we divide up the world into kinds and individuals.

However, Sidelle does not consider this to be an optional addition that we may choose to avail ourselves of once we have adopted his account. He considers this antirealism about ontology to be a consequence of his theory that cannot be avoided.⁹⁴ Sidelle claims that modality and ontology are linked to such a great extent that to be an antirealist about necessity is to also be an antirealist about kinds and individuals. For Sidelle, antirealism about modality commits us to antirealism about ontology. His motivation for this position comes from two sources: sympathy for the Lockean conception of essence, and the desire to avoid apparent pre-conventional necessity.

Sidelle claims that kinds are determined by their modal features. He cites Locke's views on essentialism, linking essences with what Sidelle (following Locke) calls the 'boundaries' of kinds. Sidelle does not specify what he means by 'boundaries' and it is by no means clear. Here I will take it to mean the membership criteria for the kind, in the hope that this is a faithful interpretation. Sidelle claims that the essence of a kind determines the boundaries of that kind, though this may not be entirely faithful to Locke, who seems to identify the boundaries with the essence:

The measure and boundary of each Sort or Species, whereby it is constituted that particular sort, and distinguished from others, is what we call its Essence.⁹⁵

Indeed, elsewhere Sidelle himself says that "essences are the boundaries of species (individuals)."⁹⁶ This aside, Sidelle states that "the ontological status of the sorts depends on that of the essences."⁹⁷ However, because Sidelle is a modal essentialist, claiming broadly that for *a* to be essentially *F* is for *a* to be necessarily *F*, if there is no pre-conventional modal reality then there is no pre-conventional essence. Sidelle says "[b]oundaries are essences, essences are necessary properties, but there is no real necessity, so no real essences, and so no real boundaries."⁹⁸ This means that:

⁹⁴ Sidelle (1989, 2008).

⁹⁵ Locke (1689), Quoted in Sidelle (1989), p.18

⁹⁶ Sidelle (1989), p.19

⁹⁷ Sidelle (1989), p.19

⁹⁸ Sidelle (1989), p.22

[I]f species (individuals) are what they are in virtue of their essences, and their essences, as such, are not ‘out there’, then Locke is pointing out that we must say that to that extent, the species (individuals) *themselves* are not, as such, ‘out there’.⁹⁹

Without essences to ground facts about what it is to be a thing (or a thing of a certain kind), there cannot be such things. If essences are conventional, then the kinds and individuals that they are the essences of must also be in some sense conventional. This is not to say that before us there were no badgers, just as Frege claims that before us there was still an equator.¹⁰⁰ Rather, the force that distinguishes these things and kinds from the rest of the world is conventional in nature and not a fact about the way the world is, independent of us.

Sidelle is also motivated by practical considerations that arise in the formulation of the GPIs. He anticipates a potential criticism that his position presupposes certain a posteriori necessities in order to give an account of necessity a posteriori. If we admit that there are kinds independent of our conventions, then one can read GPIs as saying that things are only part of a certain kind if they have certain necessary properties. Since Sidelle claims that there are no such properties, the extensions of all kind terms would then be empty.

Even when considering the application of kind terms, the conventionalist must provide an explanation of how such terms are actually applied that does not presuppose modality (or result in non-empty term extensions). It is in light of this that Sidelle develops the latter formulation of the GPI. Part of this solution is to reaffirm the Lockean claim that kinds are “the workmanship of the understanding”.¹⁰¹ Sidelle thus considers the conventionalist to be committed to antirealism about kinds and individuals (the same reasoning applied to kinds applies equally to individuals). Sidelle is not troubled by this, but rather takes it to be exactly what we should expect from such a theory.

⁹⁹ Sidelle (1989), p.19

¹⁰⁰ Frege (1884), p.35

¹⁰¹ Locke (1689), Quoted in Sidelle (1989), p.45

We are left the question of what kind of broader metaphysical picture this consequence leaves us with. Sidelle characterises the world as being made up not of kinds and individuals, but of undifferentiated *stuff*. This stuff is *pre-objectual* in that it is not divided along any metaphysical distinctions into distinct objects, but rather is the continuous material that we then categorise in accordance with our interest-relative purposes. To corrupt the metaphor of carving the joints of reality, there are no natural joints in reality, only those joints that we create as we carve it into kinds and individuals using conventions such as the GPIs. Reality is not a carcass, but a wheel of cheese. Sidelle does not think that this is an inappropriate view for the conventionalist to hold. In fact, Sidelle attributes the position to Locke:

[Locke], is claiming that while there are real features of the world out there, none of them are, considering the contribution of the world itself, essential to anything: we get essences, as such, only from our methods of carving up the world.¹⁰²

This strong link between modality and ontology is significant in the context of providing a modal account of essence. Sidelle's position provides a theoretical backing for the intuitive appeal of neo-Aristotelian accounts of essence, whereby the essence of an object or kind is significantly linked to it *being the thing that it is*, whilst also providing an account for essence in modal terms. The question remains as to whether antirealism about modality does in fact commit one to antirealism about ontology. This question is addressed in Chapter Four.

5. How Does this Give us Essence?

Sidellean conventionalism gives us an account of necessity. Necessity and possibility are inter-definable (using negation), so having an account of necessity furnishes us with an account of the basic modal notions. Modal essentialism claims that *a* is essentially *F* iff *a* is necessarily *F* if *a* exists. The modal essentialist who is also a Sidellean conventionalist can give an antirealist account of essence whereby essence is

¹⁰² Sidelle (1989), p.19

explained in terms of necessity, and necessity is grounded in convention. That is, where for a to be essentially F is for a to be F , and for $\text{nec}(Fa)$ to be the conclusion of a valid inference from Fa and an appropriate true GPI, if a exists. For example, Socrates is essentially human because Socrates is human, and being human is one of Socrates' P-properties (presumably, his species). This makes him necessarily Human if he exists, and as such makes him essentially human. The role of convention is in selecting what kinds of properties count as P-properties for the kind that Socrates is a member of (human). In conventionally selecting what the P-properties of the kind Human are we, in a sense, create that kind. We do this not in the sense that we bring something out of nothing, but in that we categorise the world into kinds by dividing things up based on their P-properties.

6. Advantages Over Other Neo-Conventionalisms

As we saw in the previous chapter, the main problem with the deflationary and Humean neo-conventionalist accounts is that they are incomplete. They give enough detail to suggest that modality is type 3 grounded in convention, but give no account as to how this should be the case. According to the deflationist we conventionally determine which worlds are the possible worlds, and according to the Human we select which kinds of facts are necessary if true. This need not be done by convention (indeed, that it is convention that plays this role is not stipulated), but regardless of whether the process by which these decisions are made is conventional or not, an explanation of that process must be given for any account of this kind to be complete. Certainly, however this process works, it will not be explicitly in terms of worlds or kinds of facts, but rather in terms of the way we think about and interact with the world around us (on a pre-theoretical level).

What sets Sidellean conventionalism apart from the other neo-conventionalist positions is that it does give an account of how this process works. This explanation is given in terms of the conventions described by the GPIs, by which we carve up the world into kinds and individuals. This is the most compelling support for Sidelle's

antirealism about objects and kinds (not his claim that without modality you cannot have them). Were it not for Sidelle's story about how we use conventions to carve our ontology out of the world, there would be no real explanation of modality being given. We would merely be making decisions about what is necessary, without giving any thought to why or how we modalise. A theory that does not address this need is no theory at all.

The stricter schematic presentation of the GPIs is phrased in terms of our resolve to apply certain terms only in specific situations (in accordance with our conventions). This plays the role of explaining the mechanism of our conventional practice, something that the other neo-conventionalisms fail to do (e.g. the deflationist claims that we conventionally select which worlds are possible, but gives no clue as to how). This provides the link between our actions and the modal facts. Furthermore, it does so in a way that is particularly congenial to a modal account of essence.

Having presented Sidelle's neo-conventionalism and shown how it surpasses the deflationary and Humean accounts in furnishing us with an antirealist account of essence, I now turn to discussing potential problems for the position.

7. The Indexicality Problem

Gillian Russell presents a problem for Sidelle's conventionalism that stems from indexicality.¹⁰³ Indexical terms are terms that change referent relative to the context of utterance. For example, 'I' when uttered by Russell refers to her, but when uttered by the author of this sentence refers to Jonathan Banks. Likewise, 'long' when ascribed to a snake may mean several feet in length, whereas ascribed to a train it may mean several hundred feet in length. The upshot of this is that sentences containing indexical terms can express different propositions in different circumstances. This is significant because the conventionalist traditionally relies on

¹⁰³ Russell (2010)

the analyticity of sentences to determine the necessity of the propositions they express.

Russell considers the Kaplan sentence ‘I am here now’.¹⁰⁴ According to Kaplan this sentence is interesting because it is seemingly analytic (no circumstances in which I utter such a sentence would be sufficient to render it false, and the meanings of the terms appear to be responsible for this), but also contingent (after all, in any circumstance in which I utter the sentence, I might not have been there, I may have been somewhere else). I will not discuss here the more basic problem indexical containing sentences pose for the traditional conventionalist by providing an apparent counterexample to the position, but rather I will focus on the ramifications Russell claims such sentences have on Sidellean conventionalism.

Russell presents a distinction from Kaplan between two different dimensions of meaning that a sentence has.¹⁰⁵ The *character* of a sentence is what is commonly understood as its meaning. It remains constant, and is what we think of when we consider a sentence independent of context. It acts as a function from the context of utterance to the proposition expressed. As such, the contribution of ‘I’ to the character of a sentence will be something like ‘the utterer of the sentence’, though it will not fix whom is being referred to unless considered relative to a specific context of utterance. The *content* of a sentence is the proposition expressed by that sentence. This kind of meaning cannot be separated from the context of utterance. It acts as a function from a circumstance of evaluation to a truth-value. As such, the contribution of ‘I’ to the content of a sentence will be to secure reference to a specific subject (e.g. me), the utterer of the sentence. In light of this distinction we can say that ‘I am here now’ is analytic in as far as it is true in virtue of its meaning, when meaning is understood as character, but not when understood as content. Indeed, sentences such as ‘I am here now’ appear not to have content, save relative to a context.

¹⁰⁴ Kaplan (1989)

¹⁰⁵ Kaplan (1989)

The sentence ‘I am here now’ presents a problem for the conventionalist. When uttered by myself it is contingently true, but when uttered by God (assuming that God exists and is necessarily both omnipresent and eternal) it is necessarily true. Call the proposition expressed by the sentence ‘I am here now’ when uttered by God Ω . There are de re and de dicto readings of the necessitation claim $\text{nec}(\Omega)$. The de dicto reading of $\text{nec}(\Omega)$ claims that necessarily, whatever the extent of space and time, God occupies all of it. The de re reading of $\text{nec}(\Omega)$ claims that for any spatiotemporal location, God necessarily occupies it. The de re reading can be taken as is, or can be read in the form of the weakened Kripke necessity, whereby for any spatiotemporal location, had that region existed it would necessarily be occupied by God. At least the de dicto and Kripke de re readings of $\text{nec}(\Omega)$ are true; this is sufficient to get the problem off the ground. Russell sets the conventionalist the challenge of accounting for the necessity of Ω .

Russell identifies that the way to disprove the Sidellean account is to demonstrate that there is no suitable analytic principle to provide the source of necessity for Ω . She gives a simplified version of Sidelle’s account to demonstrate how it cannot account for the necessity of Ω . The GPI is replaced with a simpler conditional-form principle (so that its analyticity may be more easily determined). Read ‘*HereNow*’ as an indexical spatial/temporal location predicate. Consider the following argument:

P1: If I have *HereNow* then necessarily I have *HereNow*. (Simplified version of GPI)

P2: I have *HereNow*. (Analytic Kaplan sentence)

C: Necessarily I have *HereNow*. (modus ponens from P1 and P2).¹⁰⁶

This is supposed to represent a simplified version of the reasoning employed in Sidellean conventionalism. The premise P1 replaces the relevant GPI, and P2 is the same as the required empirical claim. Whilst the argument is valid, Russell maintains that the argument is not sound because P1 is not a theorem. Any instance where ‘I’

¹⁰⁶ Russell (2009), p.14

refers to someone other than God will act as a counterexample to P1. Therefore, such an argument cannot account for the necessity of Ω .

The underlying problem that the indexicality problem illuminates is that necessity is a property of propositions, not of sentences, whilst analyticity is a property of sentences and not of propositions. If identical analytic sentences can express distinct propositions with different modal statuses, then the analytic status of the sentences cannot be the source of that modal status. However, this is not a faithful representation of the Sidelle account. By simplifying the GPI to P1, Russell excludes an element of the inference that plays a key role in securing the necessity of the proposition expressed. Consider:

GPI-HereNow: If I am a deity then if *HereNow* is my location-property, then I necessarily have *HereNow*. (GPI)

P2: I have *HereNow*. (Kaplan sentence)

P3: I am a deity. (Immodest premise)

C1: If *HereNow* is my location-property, then I necessarily have *HereNow*.
(Modus ponens from P3 and GPI-*HereNow*)

C2: Necessarily I have *HereNow*. (Modus ponens from C1 and P2)

GPI-HereNow succeeds where P1 of the previous argument failed because it (combined with P3) ensures the referent of 'I' is fixed as God. This allows for no counterexamples from the indexicality of 'I', thus enabling an explanation of the necessity of Ω .

Russell anticipates a similar response, in reply to which she says that it results in counterintuitive consequences with respect to ways that our language might have been. Russell invites us to consider a world in which there are no terms for God (perhaps a world where to have such a term would be blasphemous). In such a world there would be no such explanation for the necessity of the proposition expressed by God's utterance of 'I am here now'.

Whether or not such a reply would affect the revised argument, a response is ready to hand. This is a variation of the well-established contingency problem for conventionalism. We are invited to consider a world where our linguistic

conventions differ (not having a word for God would be an instance of this) and then confronted with the claim that the convention-determined facts at that world should differ. As we saw in Chapter One this kind of criticism is easily overcome by the conventionalist. By maintaining a convention such as

(**RIGIDITY**) What it is true to say *of* a hypothetical state of affairs, and what it is true to say *in* a hypothetical state of affairs, is to be determined by reference to our actual linguistic conventions, even if those are not the conventions that would then obtain.¹⁰⁷

the conventionalist is spared the consequences of the contingency problem. For the same reason, counterfactual variations in language should not constitute a problem for the Sidellean conventionalist.

GPI-HereNow appears to secure the necessity of the proposition expressed by ‘I am here now’ as expressed by God. However, this is merely an object language formulation of a convention, and to test it fully we must present the convention itself. How might we incorporate the content of GPI-HereNow into a conventional GPI? Here is a first attempt:

GPI (HereNow-2): If ‘*T*’ is a deity-term then if *HereNow* is the location-property of the thing denoted by ‘*T*’, then ‘*T*’ applies to something in any possible situation only if it is (has) *HereNow*.

This formulation is problematic because whereas the original GPI-HereNow could bypass the oddity of indexical terms (because it is an object language formulation), this formulation cannot. ‘*T*’ is not a deity-term (unless there is more than one term ‘*T*’, one for each agent-kind, but I shall not consider this possibility). If ‘*T*’ is not a deity-term then ‘I am here’ is not necessary when uttered by God, as the antecedent of the relevant GPI is false. However, because indexicals are something of an oddity in language, changing their meanings relative to context, it seems only fair that the

¹⁰⁷ Wright (1985), p.190 There is more detail to the response than this (see Chapter One). However, this will suffice for present purposes.

conventionalist should be given the chance to give specific rules as to how to handle sentences containing them. One might give a revised GPI. For instance:

GPI (HereNow-3): Given a context C , if ' T ' is content-synonymous with a deity-term then if *HereNow* is the location-property of the thing denoted by ' T ', then ' T ' applies to something in any possible situation only if it is (has) *HereNow*.

This formulation allows us to specify the content of the indexical term relative to the context of utterance, and from there consider content-synonymous terms (terms that refer to what ' T ' refers to in C). However, having a convention that resolves us to use ' T ' in this way in all possible situations will lead to undesirable results. ' T ' would no longer be indexical in all contexts because its content would be fixed in contexts like the one above.

GPI (HereNow-3) appears to incorporate reliance upon propositional content in determining the modal status of Ω . It represents a move away from conventionalism because the modal status is determined by the proposition itself and not the language we use to express it. This also runs the risk of assuming pre-conventional modal content because it seems to be saying something more to the effect 'if this is God then...', rather than expressing a convention governing our application of terms. This is something that Sidelle goes to great pains to avoid. GPI (HereNow-3) seems like more of a philosophical trick, rather than a legitimate convention governing how we use terms in our language. As such, we can conclude that this line of response is not a promising one (though, significantly, not for the reasons that Russell gives). There is however a simple and effective reply at hand.

8. Responding to the Indexicality Problem

By now it should be clear that trying to secure the necessity of Ω using the sentence 'I am here now' is not a promising project. However, this failure can be taken lightly by the Sidellean conventionalist. All one need do is weaken the account very slightly so as to preserve the desired results, whilst avoiding the problem. This can easily be

done. Indeed, it is not clear to me that what I am about to suggest is in fact different from Sidelle's original position. If it is then I contest it is an adjustment that should be made to avoid the indexicality problem at no apparent cost. If it is not then the indexicality problem is in fact not as problematic as it seemed. In either case I take this to be a significant result.

The Sidellean conventionalist can claim that for a proposition to be necessary is for it to be *expressible* using a sentence that is the conclusion of a valid inference from the appropriate GPIs. Just because not every sentence that expresses the proposition does this does not mean that the proposition is not necessary. The proposition expressed by 'I am here now' when uttered by God is necessary because it can also be expressed by 'God is in location *l* and time *t*' (for appropriate values of *l* and *t*). We can present the account as:

NECESSITY: A proposition $\langle \text{nec}(P) \rangle$ is true iff the proposition *P* can be expressed by an analytic sentence, or $\langle \text{nec}(P) \rangle$ can be expressed by a sentence that is the conclusion of a valid inference from a sentence expressing the claim necessitated and an appropriate true GPI.

Under this account, Ω comes out necessary, even though the source of that necessity is not at all in the sentence 'I am here now'. This has the consequence that sentences like 'I am here now', whilst expressing necessary propositions in some contexts, play no role in grounding that necessity. This does not strike me as a problematic consequence (so long as there are other sentences that can ground that necessity). The next step is to ensure that such a proposal does not have any undesired consequences, for instance, necessary propositions that we do not wish to be necessary, or propositions that we consider necessary but are not considered as such by the theory. The account is a weakening of the previous reduction, so there is no risk that propositions that come out necessary under the earlier reduction will fail to do so under the new one. The risk, if any, is that there are propositions that are not necessary under the previous account but are (undesirably) necessary under the new proposal. Possible examples might include the Quinean sentences 'the number of planets is even' or 'Giorgione was so known for his size'. These cases are not

problematic because they are examples of sentences that express more than one proposition, just like sentences containing indexicals. The sentence ‘the number of planets is even’ can express the proposition <8 is even>, or the proposition <there is an even number of planets>. The proposition <8 is even> is necessary because it can be expressed by the sentence ‘8 is even’, ‘the number of planets is even’ does nothing to affect the modal status of the proposition either way.

The indexicality problem could be revived if a case could be found whereby an indexical sentence expresses a necessary proposition, and the sentence is the sole sentence to express that proposition (or, at least, any other sentence to express the proposition is also indexical). I cannot think of such a sentence.¹⁰⁸ One may think that all of this would be much easier if we didn’t have to worry about analyticity. However, *prima facie* it seems that Sidelle’s account relies on it. In the next section I examine this reliance, and argue that analyticity need not feature in the account.

9. Reliance on Analyticity

Traditional conventionalism fell out of favour with the philosophical mainstream following the widespread repudiation of truth by convention. The primary advantage that the deflationary and Humean accounts hold over traditional conventionalism, and the characterising feature of a neo-conventionalist position, is that they do not rely on truth by convention, in particular analyticity. Two questions arise if Sidellean conventionalism relies on analyticity. First, can it really be a neo-conventionalism? Second, is it really any better than traditional conventionalism? The answers to these questions are ‘no’, and ‘not significantly’, respectively. As such, it is the task of this section to investigate, and ultimately reject, this apparent reliance on analyticity.

In the previous chapter we established three roles that convention might play in securing the truth of a sentence. These are conventional grounding types 1, 2, and 3. A sentence *P* is type 1 grounded in convention when *P* is true iff some convention *C* holds (where *C* does not feature in the meaning of *P* as in, for instance, “we hold

¹⁰⁸ And any claims as to the possibility of a counterfactual language with such a sentence can be replied to in the same way as to the contingency problem mentioned above.

convention C”). P is type 2 grounded in convention when P is true iff some other fact F holds, where C determines that the truth of P is contingent upon F. P is type 3 grounded in convention when P is true iff some fact F holds, where what counts as an instance of F is determined by C. Whilst in the previous chapter I concluded that deflationary and Humean neo-conventionalism relied on something that could be described as truth by convention, this fell under conventional grounding of types 2 and 3. However, Sidellean conventionalism explicitly relies on the analyticity of the GPIs. Analyticity, as typically understood, is a case of type 1 conventional grounding.

How can Sidelle’s reliance on analyticity be reconciled with its rejection by other neo-conventionalisms? One is left with two choices. First, one may accept Sidelle’s position at face value, and classify Sidellean conventionalism as a form of traditional conventionalism. This choice will require a defence of analyticity. Second, using the more fine-grained categories of conventional grounding introduced in Chapter One, one might argue that Sidelle’s overt reliance on truth by convention need not be understood as the problematic type 1 grounding, and that Sidellean conventionalism is, despite appearances, a neo-conventionalist theory, and not a traditional conventionalism. I take the second approach. I argue that, despite what Sidelle claims, Sidellean conventionalism need only rely on type 2 and 3 conventional grounding, and as such avoids the problems that come with reliance upon analyticity.

9.1. The Analyticity of the GPIs

Sidelle introduces the GPIs as being analytic. However, he does not do so in the standard sense that he considers them to be true in virtue of the meanings of the terms contained within them. Rather, he takes analyticity in this case as meaning that the GPIs are conventional in nature. Sidelle says:

Suppose that general principles of individuation are analytic. That is, suppose that rather than being general claims that describe features of a mind-independent modal structure of reality, these principles are instead object level

formulations of conventions we have adopted concerning how we will describe things [...].¹⁰⁹

And:

So suppose, again, that they are conventional, that is, analytic.¹¹⁰

Sidelle appears to be working with a non-standard understanding of analyticity, whereby something is analytic simply in virtue of being conventional. However, as typically understood, analytic sentences are type 1 grounded in convention, so Sidelle's use of the term 'analytic' gives the impression that the GPIs are type 1 grounded in convention. Yet this need not be the case; not everything that is conventional is analytic. For example, "Betty is driving illegally", uttered in the context where Betty is driving on the right hand side of the road in the UK, is true because Betty is driving on the right, and the law (which is conventional in nature) prohibits this. This is a case of type 3 grounding in convention because it is illegal to do things that are prohibited by the law, and convention has determined that driving on the right in the UK is prohibited by the law.

The sentence "driving on the right is prohibited by law" is true in virtue of our conventions in the mundane and uncontroversial sense that our legal conventions prohibit driving on the right). "Betty is driving illegally" is true in virtue of our conventions in that it is type 3 grounded in convention. However, neither are analytic in the traditional sense because neither are type 1 grounded in convention. We can conclude from this that when Sidelle claims that the GPIs are analytic, we need not necessarily take this to mean that they are analytic in the problematic sense of type 1 conventional grounding.

9.2. Rejecting the Analyticity of the GPIs

Sidelle insists that the GPIs are analytic, yet despite his comments on the topic it is not entirely clear what he means by this. Nor is it clear that Sidelle considers this

¹⁰⁹ Sidelle (1989), p.35

¹¹⁰ Sidelle (1989), p.36

position to deviate from standard usage, for instance he mounts a defence of analyticity against Quinean criticisms.¹¹¹ Here I claim that the GPIs need not be analytic (understood as any form of type 1 grounding), and present an account of the source of their truth. The proposed account of necessity currently takes the form:

NECESSITY: A proposition $\langle \text{nec}(P) \rangle$ is true iff the proposition P can be expressed by an analytic sentence, or $\langle \text{nec}(P) \rangle$ can be expressed by a sentence that is the conclusion of a valid inference from a sentence expressing the claim necessitated and an appropriate true GPI.

What does the account require the GPIs to be like? They must be true (as stipulated). They cannot be true in virtue of the way the world is independent of people, otherwise the theory would be a realist account of modality. The GPIs must be a result of some interest-relative practice of ours. Sidelle describes GPIs as object level formulations of conventions. It is worth distinguishing the GPIs from the conventions of which they are formulations.

The GPIs are object-level formulations of our conventional practices. The practice described by a GPI is the means by which we carve up the world. That is, by deciding what kinds of properties count as P -properties, we categorise the pre-objectual stuff of the world into kinds and objects in the way described earlier. A GPI is true in virtue of the conventions we hold, but this is not a troublesome instance of truth by convention, and it is certainly not a case of analyticity. The GPI is effectively a mere statement of a convention, just as “in the UK we drive on the left” is true in virtue of the conventions we hold about driving. We can conclude from this that for a posteriori P , the truth of $\text{nec}(P)$ is type 3 grounded in convention. $\text{Nec}(P)$ is true iff $\text{nec}(P)$ is the conclusion of a valid inference from P and an appropriate true GPI. The role that convention plays here is not in securing the truth of P or $\text{nec}(P)$, but in selecting which GPIs we use to govern our use of the terms in P . This is not truth by convention in the problematic sense, but in the mundane sense that a GPI is a description of a conventional practice. We can create an analogous account to use

¹¹¹ Sidelle (1989), Chapter 5

our driving conventions to explain the illegality of driving on the right. “Betty is driving illegally” is true iff it is the conclusion of a valid inference from Betty driving on the right hand side of the road in the UK and an appropriate true object-language formulation of our conventions (perhaps something like “for all x , if x is driving on the right hand side of the road in the UK then x is driving illegally”). “It is illegal to drive on the right” is true in the same mundane way that the GPIs are true, and “Betty is driving illegally” is type 3 grounded in convention in the same way $\text{nec}(P)$ is.

Showing that the GPIs need not be analytic nearly leaves one in a position to claim that Sidellean conventionalism doesn’t rely on analyticity at all; one task remains. The first disjunct of the theory of modality states that a proposition is necessarily true if it can be expressed by an analytic sentence. If this last vestige of reliance on analyticity can be shed, then Sidellean conventionalism can avoid the problems associated with traditional conventionalism.

9.3 Renouncing Analyticity Altogether

Having dispensed with analyticity in the GPIs, if analyticity is to be shed from the theory altogether then we must demonstrate that the theory can account for the necessity of propositions expressed by (what would count as, if we were countenancing analyticity) analytic sentences. We can absorb the first disjunct of the theory into the second. Just as with the necessary a posteriori, necessitation claims with (allegedly) analytic claims necessitated will be true iff the necessitation claim is the conclusion of a valid inference from the claim necessitated and an appropriate GPI. The challenge is to formulate GPIs that will do this job for (allegedly) analytic sentences. To show that this can be done, take the sentence ‘all bachelors are unmarried’. The GPI must be of the following form:

GPI Schema: (x) (If x belongs to kind K , then (if p is x ’s P -property, then it is necessary that x is p)).

Take this schema and substitute the kind Bachelor for x , the property Unmarried for p , and the property kind marital-status for P -property. We get the resulting GPI:

BACHELOR: If Bachelor is a marital-status kind, then if Unmarried is its marital-status, then it is necessary that Bachelor has the marital-status Unmarried.¹¹²

Thus we can form an object level formulation of the convention governing bachelors. In order to fully demonstrate how this convention works we must be able to present the GPI in the form of a conventional schema. In order for this to be done, we must be able to present the GPI in terms of our resolution to apply the term ‘bachelor’ only in certain circumstances, This can be done by taking the conventional GPI schema:

GPI Conventional Schema: If ‘ x ’ is a K-term then if p is the P-property of the thing denoted by ‘ x ’, then ‘ x ’ applies to something in any possible situation only if it is (has) p .

and making the following substitutions:

Bachelor Conventional Schema: If ‘bachelor’ is a marital-status-term, then if Unmarried is the marital-status of the things denoted by ‘bachelor’, then ‘bachelor’ applies to something in any possible situation only if it has Unmarried.

With the appropriate GPIs in place, we can present the following argument:

BACHELOR: If Bachelor is a marital-status kind, then if Unmarried is its marital-status, then it is necessary that Bachelor has the marital status Unmarried. (GPI)

P1: Bachelor is a marital-status kind. (Premise)

P2: Bachelors are unmarried. (Premise)

C1: If Unmarried is its marital-status, then it is necessary that Bachelor has the marital status Unmarried. (Modus ponens from BACHELOR and P1)

¹¹² On a first reading this GPI might appear problematic. The kind Bachelor does not have a marital-status. It’s a kind, and kinds are not the sort of things that get married. One should read ‘marital-status’ here as meaning something more like a part of what it is for Bachelor to apply to something, rather than a property of the property of being a bachelor. This works much in the same way as Zalta’s (2006) distinction between an abstract object *exemplifying* a property, and *encoding* it. As such, despite the awkward language, we can claim that Unmarried is the marital-status of Bachelor, because the bachelors are unmarried. Those who still find this problematic should be placated by the conventional schema given below.

C2: Necessarily bachelors are unmarried. (Modus ponens from P2 and C1)

It should be clear that what is going on here is no different to what Sidelle claims the other GPIs are doing. The GPI for bachelor is carving out a kind in the world where otherwise there would be none. Sidelle claims that the same is being done by the GPI for water, or for biological kinds. Cases that are commonly considered to be cases of analyticity are particularly simple examples, but the same process applies. However, we can imagine that the discovery events for the kinds involved in analytic sentences are different from those in a posteriori claims. One can imagine some person finding the first example of gold and founding the kind Gold based on this discovery (by conventionally selecting the attributes of those examples that are of importance to the kind). However, there was no occurrence where someone overturned a rock and uncovered the first bachelors. Rather, one can imagine that as language developed it became useful to have more fine-grained terms for distinguishing amongst people based on marital status, thus motivating the coining of the term and founding the (equally, if more obviously, conventional) kind bachelor. This difference is a superficial one. The important similarity in these stories is that a kind was conventionally established by selecting features that were already there (being a yellow metal in one case (say), and being unmarried in the other) and assigning them particular importance.¹¹³

One can take this further still. Using the same GPI structure we can account for logical and arithmetic necessity as well. Consider the following GPI for the logical connectives:

CONNECTIVE: (x) (If x belongs to the kind Logical Connective, then (if p is x 's truth-functional property, then it is necessary that x is p)).

In this example, p might be the truth table for the connective in question. This can be formulated as a convention like so:

¹¹³ Whilst being married/unmarried is a conventional property as opposed to the physical properties associated with gold, this is not a problem. We expect a posteriori necessities to have an empirical worldly component. We should not be surprised that necessities that are typically classified as analytic do not impose such an expectation.

CONJUNCTION: If ‘ \wedge ’ is a logical connective-term then if p is the truth-functional property of the thing denoted by ‘ \wedge ’, then ‘ \wedge ’ applies to something in any possible situation only if it is (has) p .

The fact that allows the inference to the necessitation claim (in conjunction with the GPI) is not going to be an empirical fact as in previous examples. This is not a problem so long as there is a fact of the matter that is not part of the convention in question.

It is important to note here that there is a significant difference between this application of the GPI mechanism and the one used in other, more standard examples. In other examples there has been a certain degree of choice involved. That biological origin is important to humans rather than material constitution is something that we decided based on our priorities. In Lewis’ terminology we were faced with a coordination problem and chose from multiple, equally good, coordination equilibria.¹¹⁴ Even in the bachelor case, marital status and sex were taken to be important, rather than other equally good options. In the case of the logical connectives it is not at all clear that this is the case. The connectives have very few properties as it is, and of these it is hard to think of any that are contenders for alternatives to their truth-functional properties in terms of what is important.¹¹⁵

This is part of why it is important to consider the position being developed in this project not as a strictly conventionalist one. Whilst it is true that it uses a conventionalist mechanism to achieve its end, the priorities differ from that of conventionalism. It is not my concern that the grounds of essence are in conventions strictly understood (that is, as behavioural regularities formed in response to coordination-style problems by a community that had a choice), but rather that the grounds of essence should be *in us*, in the way that we behave in and think about the

¹¹⁴ Lewis (1969)

¹¹⁵ It is worth noting that the availability of non-classical logics does not represent an alternative. Convention selects the truth-functional properties as the P-properties of the logical connectives, regardless of what those properties are. The availability of logics where the connectives have different truth-functional properties is evidence that logical connectives are not identical to their analogues in different logics. For the same reasons, the availability of alternative logical connectives does not present an alternative.

world. With this in mind it is not problematic that there was only ever going to be one option for conjunction's P- property. What is important is that that property's status as being essential has its source in the ways that we think and act, rather than the metaphysics of the people-independent world.

Paradigmatic examples of conventional distinctions come about as the results of coordination problems whereby the distinction in question is one among $n > 1$ equally viable solutions (coordination equilibria). In fact, this is built into Lewis' definition of convention. However, it is by no means clear that something's failing to be conventional in nature forces us to conclude that it is a robust feature of the world independent of people. What is important is that the distinction serves our interests, and in such a case it is not obvious that there should be multiple options available, or indeed that other forms of life would even find such a distinction to be of interest in the first place. Whilst there might be only one way to go when deciding that the essential feature of conjunction is its truth-functional properties, an intelligent species that didn't reason in terms of conjunctions would have no need to include conjunction in their logic in the first place. This reasoning does not lend further support to the claim that the logical operators are conventional in nature, but rather it diminishes the force of criticisms citing lack of alternative coordination equilibria as a problem for antirealism.

This leaves Sidellean conventionalism in the position where it does not rely on analyticity. We can present a new, more unified theory of modality in Sidellean terms like so:

NECESSITY: A proposition $\langle \text{nec}(P) \rangle$ is true iff $\langle \text{nec}(P) \rangle$ can be expressed by a sentence that is the conclusion of a valid inference from a sentence expressing the claim necessitated and an appropriate true GPI.

Without the assumed reliance on analyticity, Sidellean neo-conventionalism is in a much better position to provide an antirealist account of modal essence.

10. Conclusion

In this chapter I presented Sidelle's neo-conventionalist theory of modality. I examined it as a candidate for an antirealist modal essentialist account, and identified that the most significant challenge associated with presenting it as such is its apparent reliance on analyticity. In response to this I argued that, despite what Sidelle claims, Sidellean neo-conventionalism need not rely on analyticity, and is in fact in a position to renounce it completely. With analyticity out of the picture, the modified neo-conventionalist account is well placed to provide the foundation for a strong antirealist account of modally construed essence.

Chapter Three: Fine on Essence

1. Introduction

In this chapter I present and give commentary on Fine's neo-Aristotelian essentialism.^{116, 117} Fine's work on essence falls into two parts. First, he argues against modal essentialism, then he presents his own positive account. Fine's arguments against modal essentialism are often taken to be the strongest argument he gives in favour of his own position. The arguments take the following form. There is a base of essentialist facts that inform us regarding essence, which we have access to through intuitive understanding. This understanding comes from our pre-philosophical opinions and is expressed through natural language usage. Through this pre-theoretical understanding we gain an understanding of how essence works (much as, say, we have a pre-theoretical understanding of what counts as an instance of composition and what does not). As such, we can say that we have a well understood notion of essence, to which candidate theories must conform.¹¹⁸ According to Fine, modal essentialism establishes a group of essentialist facts that is extensionally infelicitous with the group established by our pre-theoretical notion of essence. There are essentialist claims that are true under modal essentialism that are not true according to the pre-theoretical notion. Because the pre-theoretical notion of essence cannot be adequately explained in modal terms, there is reason to believe that no modal account of essence can be found. I argue that Fine's attempts to demonstrate

¹¹⁶ This is a label I have applied to Fine's position. I know of no instance where he has applied it himself.

¹¹⁷ Fine presents this position over a series of papers. See Fine (1994), (1995a), (1995b), (1995c), (2000), (2007).

¹¹⁸ By 'well understood' here I mean that we have a confident grasp of what we take to be instances of true essential predication, just as we have a confident grasp of what we take to be instances of composition. Of course, metaphysical theories differ in what they claim to be instances of composition, and the fashion for bullet-biting in mereology is well established, but when asked if some objects compose, we have a pre-theoretical understanding that will answer 'yes' or 'no' in most normal situations.

the extensional inaccuracy of modal essentialism are not decisive, and as such Fine's stronger conclusion, that no such account could be found, is less credible.¹¹⁹

How are we to make the most of Fine's position? I propose that the debate on essentialism, as construed above, is wrong-headed. To conclude that there is such a well understood notion of essence based on our pre-philosophical opinions is premature. Instead of taking the debate to be between two rival accounts of a well understood notion, against which extensional inaccuracy is damning criticism, I propose that *i*) there may in fact not be such a well understood notion, and *ii*) that even if there is, this need not be so constraining a factor as it is often taken to be. To compare with mereology, if our pre-theoretical notion of composition were taken to be so well understood, and if extensional accuracy to this notion were held to such importance, then the theoretical landscape in the metaphysics of mereology would be very different indeed, and not, presumably, for the better.

If these claims are accepted then the debate between modal and neo-Aristotelian essentialism may be interpreted not as between competing accounts of one well understood notion, but rather as between rival notions of essence loosely corresponding to our pre-philosophical opinions. Interpreted in this way, Fine's arguments can be characterised as being definitive of a neo-Aristotelian notion of essence, rather than as critical of the modal account. This has the consequence that they do little to actively harm modal essentialism, except perhaps by presenting a competing notion of essence that is more theoretically viable or useful (e.g. by having more explanatory power). To claim that they do any more would be to beg the question against the modal essentialist. The theoretical utility of Fine's position provides much greater support under this interpretation of the debate.

To motivate this position further, in my evaluation of Fine's criticisms I identify points where interpreting the debate as I suggest is preferable to the modal essentialist (or indeed the impartial reader), as some of the subjects discussed in

¹¹⁹ An alternative, and I think successful, response to Fine comes from Nathan Wildman (2013, forthcoming). He suggests that the modal essentialist adopt a condition that only sparse properties can be essential to objects. Whilst this successfully evades Fine's criticisms, it does so by accepting a metaphysical distinction that, I take it, is similar in spirit to Fine's essentialism. The responses I present here are for those who more strictly adhere to the motivations of modal essentialism.

Fine's critique concern situations where it is unclear (to me, at least) that anything is settled by any pre-theoretical understanding of essence that we might have.

I first give a brief account of Fine's characterisation of essence. Then, in Section Two I present Fine's criticisms of modal essentialism. In Section Three I give a full exposition of Fine's position, and how he develops it. I try to present as strong a version of the theory as possible, providing further development. In Section Four I respond to Fine's criticisms in light of the developments he avails himself of in his own position, and my own proposed interpretation of the debate. I conclude that Fine's arguments fail to decisively demonstrate the extensional inaccuracy of modal essentialism, and that to see the best in Fine's work, the reader should interpret the debate on essentialism as I suggest.

1.2 Fine's Characterisation of Essence

Fine characterises essence in terms of what he calls 'real definition' rather than as something explicable in terms of any modal notion. Before presenting any theory of essence, he outlines an allegedly intuitive pre-theoretical notion of what we take essence to be. The distinction between essence and accident is better understood as parallel to that between that which features in the definition of the thing (what a thing is), and that which does not (how a thing is). The main support Fine cites for this characterisation comes from our pre-philosophical opinions, and use in natural language. In particular, Fine takes the paradigmatic form of essentialist claims (besides "*a* is essentially F") to be "*a* must be F if *it is to be the thing that it is*". Fine considers the italicised phrase to be of particular importance. He claims that this is most naturally understood as signifying that the quality attributed is definitive of that object. That is, it tells us *what* the object is, not just how it is.

It is worth noting the significance of Fine's strategy. Fine presents an allegedly intuitive understanding of essence based on our pre-philosophical opinions, and then insists that any account of essence should be able to satisfy this understanding. This effectively rules out revisionary accounts of essence. However, this understanding of

essence need not be inferred from the opinions Fine cites. Our pre-philosophical opinions about essence, and the natural language claims we use to express them, underdetermine the answers to certain important questions. The opinions behind the claim that Socrates is essentially a man, but only accidentally bearded might be accounted for in multiple ways without having to resort to the understanding of essence that Fine presents. Taking the distinction between accident and essence as parallel to that between how something is and how it must be might satisfy the distinction given in this example. There is also the question of direction of explanation for essentialist claims. However we understand essence, it seems that it is closely linked to things being the things that they are, and being of the kinds that they are. However, does an object have its essence in virtue of being the thing that it is? Or does an object's essence make it the thing that it is? Either interpretation appears to be compatible with our pre-philosophical opinions about essence. As such, it is underdetermined by ordinary usage whether the direction should be in one way or the other.¹²⁰ Fine's attempt to single out a folk notion of essence is not adequately supported by the way we think of and use essentialist claims. This jeopardises the efficacy of his argument against modal essentialism, which alleges extensional inaccuracy relative to our pre-philosophical opinions.

Fine does present his understanding of essence in another way.¹²¹ He claims that the essential properties of things are those that are of interest to metaphysicians. Essence is used to formulate metaphysical claims, and analyse metaphysical concepts. For instance, when the metaphysician theorises about personal identity, what they are really asking about is what the essence of personhood is. This, whilst perhaps a good way to elucidate the notion Fine has in mind for a philosophical audience (presumably how it is intended), is inadequate as an attempt to single out a pre-theoretical notion from the way we essentialise. Fine claims he is not denying that "anything that might reasonably be called a concept of essence" can be presented with

¹²⁰ Or indeed, whether thinking in terms of direction of explanation is wrong-headed, implying that a thing and its essence are distinct when in fact they are not.

¹²¹ Fine (1994), p.2

a modal account.¹²² Rather, Fine claims that, as far as essence is a useful notion of interest to metaphysicians, it cannot be explained in terms of modality. However, what is of interest to metaphysicians is not relevant to our pre-philosophical opinions about essence, or the way we essentialise, so it is not appropriate to reject a modal understanding of essence on these grounds. Metaphysical utility is a theoretical red herring when trying to identify a pre-theoretical folk-understanding of essence.

Fine takes his understanding of essence to be a safe assumption. He also makes the methodological assumption that extensional infelicity to this understanding of essence is a decisive objection to any rival theory of essence, thus disqualifying revisionary positions. I challenge the safety of these assumptions. I argue that our pre-philosophical essentialist opinions do not suffice in establishing a complete and well understood notion of essence, and as such there is no one standard against which candidate theories should be measured. To argue against a theory of essence because it is extensionally infelicitous with our pre-philosophical opinions is at best under-motivated and at worst question begging.

2. Against Modal Essentialism

In support of neo-Aristotelian essentialism, Fine argues against modal essentialism, the claim that essence is, or is explained in terms of, an intensional notion. Instead of accepting that the meaning of essentialist claims such as ‘Socrates is essentially a man’ can be given using possible worlds talk, Fine argues that essence is neither identical with, nor co-extensional with, an intensional notion. He claims that essence is hyper-intensional; that possible worlds talk is not fine-grained enough to express essentialist claims. There are claims that are distinct according to the essentialist that are indistinguishable in possible worlds talk. For example, the distinction between ‘being a polygon with three vertices’, and ‘being a polygon with three edges’ is hyper-intensional. Both properties pick out all and only the triangles at every possible world, so intensionally speaking there is no distinction between the two. However,

¹²² Fine (1994), p.3

one might think that each conveys different information. The former is a property about the number of vertices the polygon has, and the latter about how many edges. If one thinks that there is a legitimate difference between the two, then that difference must be hyper-intensional, because no intensional distinction is fine-grained enough to differentiate between them. Fine considers two forms of modal essentialism:

(CATEGORICAL) An object *a* is essentially F iff *a* is necessarily F.

(CONDITIONAL) An object *a* is essentially F iff *a* is necessarily F if *a* exists.

Fine's argument has two steps. He first argues that the proposed modal accounts of essence fail to produce a notion that is co-extensional with essence. To do this he gives four cases that he claims present counterexamples, where there are true necessitation claims, but no true essential predications. Fine then argues that there is reason to think that no such account could work, because of key differences between the notions of necessity and essence (which the four counterexamples reveal). Fine's reasons for the latter claim lead him to assert that rather than essence being a special case of metaphysical necessity, the converse is true, and metaphysical necessity is a special case of essence.¹²³

In arguing for the differing extensions of necessity and essence, Fine relies on what he takes to be our pre-theoretical understanding of essence. This notion is (loosely) characterised as the 'nature' of a thing, or *what it is to be that thing*. The following cases are supposed to demonstrate how the modal conception of essence conflicts with this pre-theoretical notion.

CASE 1: Socrates is both necessarily a member of his singleton set {Socrates}, and necessarily a member of {Socrates} if he exists. As such, under either modal account he qualifies as being essentially a member of {Socrates}. However, Fine, citing intuition, claims that this is false. Whilst it is necessary that Socrates \in {Socrates} (Fine does not question the necessity of the modal account, only its sufficiency), Socrates' set memberships do not tell us anything about *what it is to be* Socrates, so it

¹²³ This will be discussed further in Chapter Five.

is not essential to him that he be a member of {Socrates}, or any other set. It is worth noting in contrast that it is essential to {Socrates} that it contain Socrates as a member. This is because containing Socrates is what makes {Socrates} what it is. This displays an asymmetry in essence that is not present in necessity. It is essential to {Socrates} to contain Socrates, but not essential for Socrates to be a member of {Socrates}. This asymmetry cannot be accounted for by either of the modal accounts and, as such, essence cannot be co-extensional with the modal notions they employ.

CASE 2: All true identity and non-identity claims are necessarily true.¹²⁴ Socrates is distinct from the Eiffel tower; he is therefore necessarily distinct from it. However, Fine rejects the claim that Socrates is essentially distinct from the Eiffel Tower, “for there is nothing in his nature that connects him in any special way to it”.¹²⁵ Again the pre-theoretical understanding of essence that Fine is trying to promote compels one to only accept an essentialist predication when the predication says something about what it is to be the thing of which it is being predicated. Whilst it may be necessary that Socrates is distinct from the Eiffel Tower, this doesn’t tell us anything about the nature of Socrates. As such, Fine sees no reason why the Eiffel Tower should feature in Socrates’ essence.

CASE 3: Consider any claim necessitated, for example that $2+2=4$, or that there are infinitely many prime numbers. These claims are necessary if true. Thus, for either, “it is necessarily the case that this truth should hold if Socrates exists.”¹²⁶ As such, under modal essentialism, it is essential to Socrates that $2+2=4$. However, that $2+2=4$ has nothing to do with what Socrates is, so we should resist the claim that it is essential to him. Furthermore, as Fine agrees that true essential predications are also

¹²⁴ Here and throughout, when talking of identity and non-identity claims I intend them to hold between things as referred to rigidly. This excludes from present consideration contingent identity and non-identity claims such as ‘Socrates is the teacher of Plato’ and ‘Socrates is not the founder of the Academy’, both of which are true, but might have been false.

¹²⁵ Fine (1994), p.4

¹²⁶ Fine (1994), p.5

necessary, the essence of any object will (according to either modal account) include the entirety of the essence of every other object.

One might resist Case 3 by pointing out that it does not fit the characterisation of modal essentialism that was presented earlier. Recall, CONDITIONAL claimed that a is essentially F iff a is necessarily F if it exists. An important part of this is that it specifically concerns predications. The proposition $\langle \text{Necessarily } 2+2=4 \text{ if Socrates exists} \rangle$ does not take this predicational form, so one might want to reject it out of hand. One must rephrase this claim to take the form of a predication if it is to be relevant. One can say that Socrates necessarily *is such that* $2+2=4$ (or, indeed, such that the content of any claim necessitated is the case), but he is not essentially such that $2+2=4$ (etc.). By lambda abstraction one can reach the predicate $[(\lambda x)((\forall y)(y = x))]$, a , which could be used to express claims necessitated such as the claim that a is such that everything is self-identical. However, it is worth noting that it is not obvious how one might use this strategy to express the predicate ‘is such that $2+2=4$ ’, or that there are infinitely many prime numbers, without introducing sentence letters into the language. One may be sceptical about such properties,¹²⁷ but for the moment I will take this to be a satisfactory presentation of the problem.

CASE 4: According to CONDITIONAL, Socrates essentially exists. Trivially, it is necessary that Socrates exist if he exists. However, Fine rejects the claim that Socrates essentially exists, claiming that it is absurd. Fine is willing to countenance things that exist essentially, calling such things ‘essential beings’, but reserves such status for the likes of God. Even Socrates is excluded from this elite group.

Fine identifies the common factor in these cases as the lack of asymmetry in the necessary connections cited, and which is present in the corresponding essentialist connections. This asymmetry makes essence more fine-grained than necessity, rendering necessity unable to provide an explanation of essence.

¹²⁷ Indeed, Wildman’s response adds a condition to the modal account that only sparse properties may be included in the essence of an object.

Fine thinks that there is further intuitive reason to reject even the prospect of a modal account. When considering natural language instances of essentialist claims, Fine claims that they typically take the form ‘*x* must be *F* if it is to be the thing that it is’. There is no provision made by the modal accounts for what the qualifying clause ‘if it is to be the thing that it is’ might mean. Under the modal accounts it appears to be either redundant, or vacuous. Fine claims that the characteristic qualifying clause of essential predications seems to be better accounted for if it indicates that the essence is acting as a definition of the object in question.

It is worth noting that Fine’s strategy, by relying on our pre-philosophical opinions about essence, depends on denying certain essentialist claims (that Socrates is essentially a member of {Socrates}, for example) on the grounds (to put it somewhat unflatteringly) that they sound odd. However, there is a significant difference between what is odd or say and what is false. In defence against Fine’s arguments, the modal essentialist can appeal to the distinction between what is sensible or germane to say in everyday discourse, and what is true by our metaphysical theory. In my discussion of and response to Fine’s criticisms I suggest that, since our pre-philosophical opinions do not fully determine a complete and well understood notion of essence, a certain amount of revisionary consequences should be acceptable in a theory (or at the very least, should not be a damning criticism in its own right). When the results of a theory of essence sound odd or contrast with our pre-theoretical understanding, we should bear in mind the distinction between what merely sounds odd, and what is really false.

3. Theory of Essence

3.1 Characterisation of Essence as Definition

The reader may have noticed that Fine appears to consider essence in two ways. At times he discusses essence in terms of true propositions concerning an object, and at others he refers to objects having properties essentially. He apparently leaves the essentialist to decide whether they want to frame the theory in terms of the property-

based characterisation, or the proposition-based characterisation. However, the question that is seemingly left unanswered is the metaphysical question concerning what essentialist claims are actually about. Here I remain neutral concerning which characterisation to take, and the metaphysical ramifications that come with them.

According to Fine, essences act as the ontological equivalent of linguistic definitions. Just as we can provide linguistic definitions for terms, stating what those terms mean, essences act as the *real definitions* of objects, stating what those objects are. With this in mind, Fine presents an extended comparison between essence and analyticity. He claims that there is “a systematic analogy between necessity and analyticity, on the one hand, and essence and meaning, on the other; as essence is to necessity, so is meaning to analyticity.”¹²⁸

Fine gives a ‘traditional explication’ of analyticity. The meaning of a term is identified with a set of defining sentences, and the analytic sentences are those that are logical consequences of the totality of definitions.¹²⁹ “[A] given analytic statement is derived from definitions which in a significant sense provide one with the meanings of the individual terms.”¹³⁰ From this understanding of analyticity, Fine claims, it is reasonable to identify a relativized form of analyticity, whereby a sentence is true in virtue of the meaning (i.e. a logical consequence of the set of defining sentences) of some words, and not others. For example, ‘bachelors are unmarried men’ is taken to be analytic of ‘bachelors’ and not of ‘unmarried’ or ‘men’. ‘Bachelors are unmarried men’ is true in virtue of the meaning of ‘bachelor’ alone, whereas neither ‘unmarried’ nor ‘men’ alone are sufficient to render ‘bachelors are unmarried men’ true.

As indicated above, Fine discusses two ways of thinking about essence. On the one hand he characterises essence in terms of properties that objects have, and on the other in terms of propositions made true by objects. Fine suggests that we identify an object’s essence with the class of its essential properties, or alternatively, with the class

¹²⁸ Fine (1994), p.8

¹²⁹ Fine (1994), p.8

¹³⁰ Fine (1994), p.11

of propositions that are true in virtue of the nature of that object.¹³¹ Here I focus on the propositional form. I do this for two reasons. First, in Fine's discussion of essence there is frequent mention of properties that are somewhat controversial. Properties like '... is a man or a mountain' and '... is such that $1+1=2$ ' play a role in Fine's work on essence, yet it is not clear that there are such properties (even once one gets past any concerns about the existence of properties *tout court*). However, the corresponding propositions are no more controversial than the existence of any proposition. <Socrates is a man or a mountain> and <Socrates is such that $1+1=2$ > are not controversial in the same way.¹³² Second, as will become clear, the nature of the 'true in virtue of the nature of' locution, and the role it plays in the theory of essence, is much clearer when thinking of essence in terms of propositions. This is not to say that it is otherwise unintelligible – indeed, Fine seems to place no importance in the choice between the property based and propositional characterisations – but adopting the propositional form is for now a prudent methodological move. The purpose of this discussion is not to examine this controversial formulation, but rather to further develop the position along the lines of the Finean project. I will trust that whatever is said of the propositional form can be adjusted so as to apply the property form.

Essence construed as real definition works in the same way as the definition of a term. The definition of a term is the set of definitional truths that concern that term. Fine takes these definitional truths to be true in virtue of the meanings of their defined terms.¹³³ As such, the definition¹³⁴ of a term is the set of propositions that are true in virtue of that term meaning what it does. Sentences are analytic of that term when they are logical consequences of its definition. Parallel to this, we can say that the essence of an object is the set of propositions that are true in virtue of that

¹³¹ Fine (1995a), p.55, (1995b), p.275

¹³² One may think that <Socrates is such that $1+1=2$ > is controversial, but at least it is less so than the corresponding property.

¹³³ Fine (1994), p.11

¹³⁴ Fine in fact uses the term 'meaning' here, but from the usage elsewhere (and in order to avoid the uncomfortable claim that the meaning of a term is identical to the set of propositions that are true in virtue of its meaning) I substitute in 'definition'. I do not see that anything is lost by this, and believe it is more faithful to Fine's intentions.

object being what it is (the ‘identity’ or ‘nature’ of that object). Furthermore, as a statement that is analytic in the relativized sense is true in virtue of the meanings of (some of) the terms contained within it, a necessary proposition is true in virtue of the essences of (some of) the objects contained within it.¹³⁵

It is important to note that Fine does not take this account to constitute an analysis of essence. He merely proposes a model by which we can understand how essence works, whilst taking it as primitive. However, *prima facie* one might think there is cause to think that this is under-ambitious of Fine. The notion of truth in virtue of identity, where what ‘identity’ means is left somewhat mysterious (Fine uses ‘identity’ and ‘nature’ interchangeably), suggests that essence is being reduced to some metaphysical primitive: identity/nature. To ease this ambiguity, I will use ‘nature’ in the place of ‘identity’. As Fine admits, “[t]he idea of what something is, its *identity* or *being*, is notoriously obscure”,¹³⁶ and reducing essence to such a mysterious primitive would make for little progress in the debate. Indeed, this is not what Fine means. To clarify the issue further it is important to specify just what Fine means by his use of the ‘... is true in virtue of the nature of ...’ locution.

3.2 True in Virtue of the Nature of...

Fine does not take the ‘... is true in virtue of the nature of ...’ locution as a sign that there is some other thing, the object’s ‘nature’ upon which the truth of a proposition is dependent. Indeed, he denies that ‘the nature of ...’ is a significant grammatical component of the locution.¹³⁷ Fine takes the locution to signify an unanalysed relation between a proposition and an object.¹³⁸ Call this relation TIVON.

Whilst no analysis of TIVON can be made, some remarks can be made to further clarify when it holds. A proposition P bears TIVON to an object (or group of objects) when that proposition is true because those objects are the objects that they

¹³⁵ These are the basic lines along which Fine develops his essentialist theory of modality.

¹³⁶ Fine (1995b), p.270

¹³⁷ Fine (1995a), p.54, fn.2 Fine in fact mentions ‘the identity of ...’, but uses ‘identity’ and ‘nature’ interchangeably.

¹³⁸ Fine (1995b), p.273

are. Note that this is not intended to be the same as those objects' mere existence. For instance, $\langle \text{Socrates is human} \rangle$ bears TIVON to Socrates, but $\langle 1+1=2 \rangle$ does not, even though it is true whenever Socrates exists. Likewise, $\langle \text{Socrates is bearded} \rangle$ does not bear TIVON to Socrates, even though it is made true by his existence. As such, TIVON is a hyper-intensional notion. This aside, the important point is that we have identified the key primitive of Fine's account, the 'true in virtue of the nature of' relation.¹³⁹

3.3 Predicational and Sentential Forms

Following the above characterisation of the source of essentialist claims, Fine presents two forms that essentialist claims can take: predicational, and sentential. Under the predicational form, take an essentialist operator ' ϵ ' to modify predicates to form complex essentialist predicates of the form ' ϵF ' (... is essentially-F).¹⁴⁰ Under the sentential form, take an essentialist operator ' ϵ ' to modify sentences to form essentialist claims of the form $\epsilon_a(Fa)$ (it is true in virtue of the nature of a that a is F; i.e. that $\langle Fa \rangle$ bears TIVON to a).¹⁴¹

Whilst Fine wants to retain both forms of expression, he stresses that they are not mere notational variants, as the predicational form allows for distinctions that the sentential form does not. For example, when considering the essential self-identity of a , the predicational form allows one to distinguish the claim $[\epsilon \lambda x(x=x)],a$ from $[\epsilon \lambda x(a=x)],a$ or $[\epsilon \lambda x(x=a)],a$ or $[\epsilon \lambda x(a=a)],a$.¹⁴² The sentential form on the other

¹³⁹ An interesting, and potentially problematic, issue arises from taking TIVON as primitive. It is not clear what relation, if any, TIVON bears to truth simpliciter. One's first thought is to claim that the former is a special case of the latter, but if TIVON is taken as primitive then this cannot be the case, nor can any other explanation of what it is that being true in virtue of the nature of something, and simply being true have in common. Fine says nothing on this matter

¹⁴⁰ My notation differs from that used by Fine, who favours the ' \Box ' symbol. I use ' ϵ ' so as to avoid ambiguity between essence and necessity.

¹⁴¹ Despite the mention of truth in the sentential modifier, this should not be understood as part of a meta-language, but as a primitive operator within the object language.

¹⁴² Where $a = \text{Socrates}$: $[\epsilon \lambda x(x=x)],a$ means that Socrates is essentially such that he is self-identical; $[\epsilon \lambda x(a=x)],a$ means that Socrates is essentially such that Socrates is identical to him; $[\epsilon \lambda x(x=a)],a$ means that Socrates is essentially such that he is identical to Socrates; $[\epsilon \lambda x(a=a)],a$ means that Socrates is essentially such that Socrates is identical to Socrates.

hand can only express the claim $\varepsilon_a(a=a)$. Fine is inclined to regard the predicational form as fundamental.

3.4 The Extent of Essence

Having described the source and form of essentialist claims, Fine turns his attention to considering the extent of essentialist claims (i.e. what kind of claims fall within the essence of an object).

3.4.1 Consequential/Constitutive Essence

Fine asks if essential predication is closed under logical consequence.¹⁴³ That is, when Socrates' being essentially F logically implies that Socrates is G, is Socrates essentially G? If the answer to this question is 'yes' then the Finean essentialist encounters the same problem as the modal essentialist, in that the Finean essentialist will be forced to include propositions in the essences of objects that our pre-philosophical opinions insist we exclude. In particular, three forms of proposition are logical consequences of Socrates being F.

- (1) <Socrates is F or H, where H is some arbitrary predicate>
- (2) <Socrates is F or not F>
- (3) <Socrates is F and such that I, where I is an arbitrary claim necessitated>¹⁴⁴

Fine criticises the modal account of essence on the grounds that it is extensionally inaccurate; it includes propositions in the essences of objects that don't say anything about what those objects are. Cases like 1, 2, and 3 are problematic for just the same reasons. If Fine cannot exclude them from the essences of objects then he will be subject to a *tu quoque* response from the proponents of modal essentialism. For this reason, Fine must resist logical closure. To do so, Fine distinguishes between constitutive and consequential conceptions of essence.

¹⁴³ Fine (1995a)

¹⁴⁴ This is by no means an exhaustive list, but it is sufficient for the points made in this discussion.

It is **constitutive** of *a*'s essence that *a* is F iff *a* is essentially F, and that *a* is F is not a logical consequence of *a* being essentially ϕ (where ϕ is any 'more basic' part of *a*'s essence).

It is **consequential** of *a*'s essence that *a* is F iff *a* is essentially F, and that *a* is F is a logical consequence of *a* being essentially ϕ (where ϕ is any 'more basic' part of *a*'s essence).¹⁴⁵

The constitutive essence is taken to be 'directly definitive of the object'. It is this, and not the consequential essence, that is of interest and importance to the metaphysician. By adopting the constitutive, and rejecting the consequential, conception of essence the problematic closure propositions do not enter into objects' essences. This allows Fine to maintain the definitive characterisation of essence that he is keen to promote.

Fine recognises a significant concern with the distinction. It is difficult to know both where the boundary between constitutive and consequential essence should lie, and how it should be decided upon. In order to maintain an intelligible characterisation of the constitutive conception of essence, we need an understanding of what a 'more basic' part of an object's essence is. Without any apparent way to draw this line using tools already available to him, Fine has little choice but to either introduce more ideology to make the distinction, or to abandon constitutive essence. Thus, short of finding a way to draw such a distinction safely, Fine suggests that we retain the consequential conception of essence at the expense of the constitutive.¹⁴⁶

However, Fine claims that there is a way to retain the benefits of the constitutive conception whilst retaining the conceptual economy of the consequential conception. He says that we can generalise away problematic propositions that are gained by logical closure. It is a part of Socrates' consequential essence to be F or not F. However, it is a part of every other object's consequential essence as well. As such, it can be effectively excluded from consideration because it can be generalised out.

¹⁴⁵ Fine (1995a), p.57

¹⁴⁶ Presumably Fine doesn't intend to keep the consequential conception of essence as is, as the definition he gives also mentions 'more basic' parts. I will not press the matter here as it seems safe to assume that an alternative formulation could be found.

Likewise, if the generalisation of a proposition is contained in the consequential essence of an object, then both the proposition and its generalisation can also be generalised out. Fine claims that this is a non-ad hoc way of excluding problematic propositions from objects' essences. The claim is that the essential propositions of interest are those that cannot be generalised away. We can present the principle like so:

GENERAL: Any proposition that is part of the consequential essence of an object can be generalised away either if it is part of the essence of every object, or if the generalisation of that proposition is also contained in the consequential essence of the object.¹⁴⁷

An example of the first disjunct: everything is such that it is F or not F, and as such being F or not F can be generalised out. An example of the second disjunct: the generalisation of $\langle \text{Socrates} = \text{Socrates} \rangle$ is $\langle (x)(x = x) \rangle$. $\langle \text{Socrates} = \text{Socrates} \rangle$ is necessarily true, and as such will feature in the essence of any object a , because it is a logical consequence of any proposition in a 's essence. However, $\langle (x)(x = x) \rangle$ is contained in the consequential essence of a (because it too is metaphysically necessary), So both it and any specific instance of it can be generalised away.

This kind of generalisation response is in the spirit of Aristotelian essentialism as traditionally understood, for which it is characteristic of x being essentially F not only that x is necessarily F, but that there be some $y \neq x$ such that y is not necessarily F. As such, it seems only proper that one should exclude properties that are universally instantiated from being essential to objects.

Whilst successful in excluding 2 ($\langle \text{Socrates is F or not F} \rangle$), the generalisation strategy does nothing to exclude 1 ($\langle \text{Socrates is F or H, where H is some arbitrary predicate} \rangle$) or 3 ($\langle \text{Socrates is F and such that I, where I is an arbitrary claim necessitated} \rangle$) from Socrates' essence, as neither feature in the essences of everything (assuming that F is something non-universal, such as 'is human'), nor do

¹⁴⁷ Fine (1995a), p.59

generalisations of 1 or 3 feature universally in the essences of everything.¹⁴⁸ The question remains as to how to go about excluding 1 and 3.

3.4.2 The Role of Logical Concepts

Fine himself has developed the resources for overcoming this problem, specifically his account of logical concepts.¹⁴⁹ Fine claims that logical concepts have essences, and that these are the source of logical necessities. He is committed to the essences of logical concepts in as far as he uses them in his essentialist theory of necessity. Fine does not say anything about what kinds of things these logical concepts are, but since logical concepts have essences, they must be viable relata for the TIVON relation.

If it is the case that the TIVON relation can hold between propositions and logical concepts, logical concepts must be taken into consideration in the above problem. Because 1 and 3 are logical consequences of Socrates being F, even if essence is closed under logical consequence 1 and 3 do not bear TIVON to Socrates alone, but rather to plurals of Socrates and certain logical concepts, for example, the plural of Socrates and disjunction (in the case of 1), or of Socrates, conjunction, and whatever it is that I is true in virtue of (in the case of 3). We can use this to justify excluding any proposition that does not bear TIVON *solely* to Socrates. This would mean that 1 and 3 are essential to the plurals, but not to Socrates, or any of their other proper parts, alone. Fine is already committed to reifying logical concepts (at least to the extent that they can be relata for TIVON), so the ontological cost (for Fine at least) is minimal.

This reveals a potential concern. If the logical consequences do not bear TIVON to Socrates alone, but to Socrates and the relevant logical concepts, then is it not the case that such a plurality must also include the concept of logical consequence

¹⁴⁸ Unless one is willing to include second order generalisations, where the predicates are replaced with predicate variables. Suffice to say, such a response would be far too strong, excluding much more than desirable.

¹⁴⁹ Fine says nothing to elaborate on exactly what logical concepts are. For now I accept this at face value and will not challenge Fine on it. However, in Chapter Five I look closer at exactly what this means, and whether it is problematic.

itself?¹⁵⁰ Otherwise, what will take one from the essences of Socrates and disjunction to 1? Furthermore, what of Socrates' essence in the first place? Socrates' essence is made up of the propositions that bear TIVON to him. TIVON is not the same as logical consequence. However, in the analogy between essence and analyticity, TIVON plays the same role in real definition as logical consequence does in nominal definition. Would it not be the case that there are no propositions essential to Socrates alone, but rather only propositions essential to the plural of Socrates and the TIVON relation? The response looks in danger of excluding too much.

In order to show why this is not the case, first we must concern ourselves with the essences of logical concepts.¹⁵¹ Fine suggests that we should take the essences of logical concepts to be inferential in character, rather than propositional. The essences of logical concepts like disjunction are not made up of propositions of the form <if P then P or Q>, but rather they are made up of inferences that the concept licences, such as the inference from P to P or Q.^{152, 153} For example, one might take the essence of the logical concept corresponding to one of the logical connectives to be made up of the introduction and elimination rules for that connective. As Correia points out, there are some features of a logical concept's essence that are propositional in nature, for instance that it is a concept. Therefore we should distinguish between the properly logical essence of a logical concept, which is inferential in nature, and the non-properly logical essence, which is propositional in nature.¹⁵⁴

Fine characterises the essences of logical concepts as inferential for reasons different to those that are relevant here.¹⁵⁵ However, securing the inferential character of the essences of logical concepts removes the need to posit logical

¹⁵⁰ One might think that instead of jumping straight to logical consequence we could include some conditional rule like modus ponens. However, going down such a road would leave us susceptible to a Carroll (1895) style regress.

¹⁵¹ Whatever logical concepts may be, Fine is committed to them and their having essences, and so in developing Fine's position it seems fair to employ them.

¹⁵² Fine (1995a), p.58

¹⁵³ Exactly what inferences are supposed to be is not made clear, but Fine allows himself them, so I shall as well. Nor will I consider the potential problems that may arise from this move. For example, the inferential nature of the essences of logical concepts implies that the TIVON relation must also be a relation between inferences and concepts.

¹⁵⁴ Correia (2012), p.8

¹⁵⁵ See Fine (1995a), p.57

consequence as a concept in its own right. Rather than having the essence of Socrates and the essence of disjunction, and then needing something in addition to get from those to a claim like 1, the inclusion of disjunction licences the inference itself (by the inferential character of its essence). In effect, logical consequence is ‘built into’ the logical concepts. As such, 1 bears TIVON to Socrates and disjunction, rather than Socrates, disjunction, and logical consequence. I contend that if this is the case for logical consequence, then the same can be the case for the TIVON relation.

One might think that because the TIVON relation is metaphysical in nature, not logical, it acts differently to logical consequence. Whatever it is about the TIVON relation that links objects to propositions, it is not logical consequence. However, TIVON is modelled on logical consequence; doing a metaphysical analogue of the task logical consequence does in definition. The proposal is akin to claiming that, just as the notion of logical consequence is built into the logical concepts (via the inferences that they license), so too TIVON is built into everything, such that certain propositions will be true, simply because an object is the object that it is. This fits with the initial characterisation given of TIVON above. If this is accepted then the problem of excluding 1, 2, and 3 is resolved. 1 – 3 will not be essential to Socrates because they do not bear TIVON to Socrates alone, but rather to plurals containing Socrates. This response does not go too far, because (now that TIVON is ‘built in’ just as logical consequence is) propositions like <Socrates is human> are essential to Socrates alone.

3.4.3 Reviving Constitutive Essence

Using this strategy to merely save Fine’s jury rigging of consequential essence is under-ambitious. Understanding logical consequence and the TIVON relation in this way allows us to properly demarcate the constitutive conception of essence. One can take the constitutive essence of an object to be the class of propositions that bear TIVON to that object alone. The consequential essence of that object will be the class of propositions that bear TIVON to the pluralities that the object is a part of.

Defining constitutive essence in this way provides the Finean essentialist with a way to deny that essence is closed under logical consequence. It does not require the introduction of any new ideology that the Finean is not already committed to. It ensures that constitutive essence meets the Finean criteria set out in the introduction to this chapter. The primitive TIVON relation is well suited to this role, ensuring that essence is co-extensive with our (alleged) pre-theoretical understanding of essence.

3.4.4 Mediate/Immediate Essence

The second question concerning just what should be included in the essence of an object asks whether the essence of an object should include the essences of the objects that feature within that object's essence. Fine distinguishes between the immediate essence of an object, those essential properties that the object has and that are definitive of it, and the mediate essence of an object, those essential properties of the objects that feature in its essence. For example, $\{\{\text{Socrates}\}\}$ essentially contains $\{\text{Socrates}\}$, which in turn essentially contains Socrates. $\{\{\text{Socrates}\}\}$ immediately essentially contains $\{\text{Socrates}\}$, but it is only mediately essential to $\{\{\text{Socrates}\}\}$ that $\{\text{Socrates}\}$ contain Socrates. The question is this: when we are considering the essence of an object, do we have immediate or mediate essence in mind?

The mediate notion of essence can be defined in terms of the immediate, but not vice versa. The mediate conception of essence can be defined in terms of the immediate by ensuring that the immediate essence of an object is closed under the following rule: The essence Γ of object x (where Γ is the set of propositions that are true in virtue of x 's nature) will contain the essence Δ of object y whenever y is mentioned in a member of Γ . No similar rule can be found to define immediate essence in terms of the mediate. This is because agreement on the facts about the mediate essence of an object does not secure the facts about the immediate essence of that object. For example if it is essential to x that it came from egg y , and essential to y that it be a chicken egg, then it is mediately essential to x that y be a chicken egg. However, it may also be the case that it is immediately essential to x that y be a

chicken egg. After all, is it not definitive of x that it came from a chicken egg? Regardless of whether it is a part of both x 's mediate and immediate essence is besides the point. What is of importance is that this is undecided by the facts about the mediate essence.

Fine describes such disagreement about the immediate essence of an object once the mediate essence is established as rare or inconsequential. As such he claims that "we can, for the most part, confine our attention to what belongs to the mediate essence of things and let considerations of what belongs to the immediate essence ride on the result."¹⁵⁶

3.5 Possible Sources of Regress

Whilst Section 3.4 develops Fine's discussion of the extent of essentialist claims, in Section 3.5 I present Fine's discussion of two possible sources of regress that come from the way he develops his theory of essence.

3.5.1 Reflexive Essence

Fine distinguishes between an object having a self-relating property and having a reflexive property. A property of an object is self-related if that property involves the object itself (e.g. a 's property 'being identical to a '). Corresponding to each self-relating property is a reflexive property (e.g. being self-identical). Fine asks what the connection is between essentially having one and essentially having the other.

Fine claims that if an object essentially has a self-related property then it must also essentially have the corresponding reflexive property, but that essentially having a reflexive property does not require that the object essentially has the corresponding self-related property. When a relation features in the essence of an object, whatever the object is linked to by the relation also features in the object's essence. An object that essentially has a self-related property features in its own essence. Since Fine

¹⁵⁶ Fine (1995a), p.62

claims that an object is dependent not only on the existence of those objects that feature in its essence, but on their essence as well, when an object features in its own essence its identity is presupposed in its essence. This results in circularity unless the inclusion of the self-relating property can be attributed to the inclusion of its reflexive counterpart. The reflexive property gives no mention of any object, and thus its inclusion in the essence of an object will not result in any such circularity. Essentially having the corresponding reflexive property provides a basis by which the object can be non-circularly said to essentially have the self-related property. However, there is no equivalent reasoning to compel one to include a self-relating property in an object's essence, unless it is as a consequence of the inclusion of the reflexive property.

It is worth noting that the above considerations are only relevant to the predicational formulations of essentialist claims. If one only formulates essentialist claims in the sentential form, there can be no distinguishing between the self-relating and reflexive properties. Socrates' essentially being identical to Socrates, and Socrates essentially being self-identical are both expressed as $\varepsilon_a(a=a)$ under the sentential form. Further to this is the relevance this has to the property-based and proposition-based conceptions of essence. If one adopts the property-based conception and the predicational formulation, then there is a distinction, whereas if one accepts the proposition-based conception and the sentential formulation, then there is no distinction to be made.

3.5.2 Reciprocal Essence

Fine identifies what he considers to be a second, indirect, potential source of circularity. Whereas the threat from self-related properties came through an object ineliminably containing itself in its essence, the threat of this section comes from the possibility of two objects ineliminably containing each other in their essences. Fine gives two examples where he takes this threat to be manifest.

The first example concerns fictional characters. The characters Jeeves and Wooster are defined in terms of their relationship to each other, namely, that one is

valet to the other. As such, it should be essential to each that Jeeves is the valet to Wooster. Each features in the essence of the other, and each in turn contains the other in their essence, leading (via an indirect route) to the kind of circularity avoided previously, where each object features in its own essence.

The second, and perhaps more metaphysically relevant example concerns points in Euclidean space, which Fine takes to be defined in terms of their spatial relations to each other. Consider two points x and y in Euclidian space. Each is essentially distinct from the other (because their definitions are given in terms of being distinct from every other point in space), and thus each features in the other's essence, whilst also having the other featured in its own essence. It is worth noting that this issue is relevant for both the predicational and sentential formulations of essentialist claims. The relevant claims can be formed under both formulations.

Fine claims that this problem can be avoided. Instead of claiming that it is essential to Jeeves that he is Wooster's valet, and that it is essential to Wooster that he has Jeeves as a valet, one should say that it is essential to both Jeeves and Wooster, together, that the former is the valet to the latter. "The apparently circular understanding of each, taken separately, is replaced by a simultaneous understanding of the two taken together."¹⁵⁷

Fine takes it to be reasonable to forbid reciprocal relations from featuring in the constitutive essences of objects on the grounds of avoiding circularity. He asks whether the same should be the case for consequential essence, or if such relations should be allowed. Fine thinks that the essentialist can go either way on this, but that how one answers the question will determine whether dependence is a one-way or a two-way relation.

3.6 Essential Manifold

The final distinction Fine makes in developing his theory of essence stems from the possibility of objects having multiple definitions. It has been presupposed that each

¹⁵⁷ Fine (1995a), p.65

object has a single unique essence or definition. In calling this assumption into question Fine tries to further clarify the relationship between essence and definition.

To motivate these considerations Fine discusses the Aristotelian view of colour. Under the Aristotelian view, colours are defined in terms of their instances. Red can be defined as ‘the colour of x ’, where x is some red object. This means that there is a separate definition of red for each and every one of the red things. Not only are these definitions distinct in that they have different contents, but Fine contends that they are not plausibly parts of a larger, more comprehensive definition, as (on the assumption that the Aristotelian view is acceptable) the addition of more exemplars does nothing to improve the definition.¹⁵⁸

Fine takes this to be sufficient motivation for the discussion, and distinguishes between three notions of essence corresponding to different definitions: the *manifold*, *component*, and *common* essences. A component essence corresponds to a single definition for a thing. The manifold essence is the range of the possible definitions, each of which has a corresponding component essence. The common essence is that which is common to all of the component essences. The manifold essence is unique “but multifarious rather than propositional in character.”¹⁵⁹ The component essences are not unique, but are propositional in character, and the common essence is both unique and propositional. It is worth noting that whilst the common essence of a thing is necessary to it, the component essences are not. As demonstrated in the case of an Aristotelian theory of colour, red need not be defined as the colour of telephone boxes or buses, so it need not have all of its component essences.

If an object has a property in virtue of some definition of it, then the object has that property *definitively*. If the object has a property in virtue of every definition of it, then the object has that property essentially. The common essence of an object is necessary to it, whereas the component essence need not be necessary. The definitive properties of an object are not determined by its essential properties. This has the important consequence for metaphysics that in order to give a full account of the

¹⁵⁸ Fine (1995a), p.67

¹⁵⁹ Fine (1995a), p.67

nature of something, one must give both the definitive and essential properties of that thing.

4. Responding to Fine

We are now in a position to respond (or, at the very least, present potential replies that warrant discussion) to each of the four cases presented in Section Two. These replies do not vindicate the modal essentialist, but are at least successful in calling into question the decisiveness of Fine's critique.

4.1 Case 3

Case 3 claims that because every claim necessitated is necessary, Socrates will be necessarily such that those claims are true. According to modal essentialism this means that it must be essential to Socrates that (for example) $2+2=4$, or that the Eifel Tower is a tower. Fine denies that such claims have anything to do with the essence of Socrates, and so modal essentialism is extensionally inaccurate. A first attempt at a reply to Case 3 might be to assert that claims necessitated such as the claim that $2+2=4$ can be generalised away from the essences of objects in the way Fine allows himself. Everything is such that $2+2=4$, and as such, it can be safely excluded. However, this response may be too strong, excluding more than is desirable. It is the case that x (and indeed everything) is essentially such that a is essentially F , because, as Fine says, essentialist claims are themselves necessary.¹⁶⁰ Since everything essentially has the property of being such that a is essentially F , it is generalised away in the fashion that Fine allows. However, this ought to be resisted because it is still important to a that it be essentially F .

This problem can be avoided. Whilst 'being such that a is essentially F ' is generalised away, 'being essentially F ' is not generalised away from a . As such, a remains essentially F , but nothing else is essentially such that a is essentially F . This

¹⁶⁰ Fine (1994), p.5

can be better elucidated using the predication form of essential expression that Fine uses.¹⁶¹ Whilst it is true that $\forall x([\varepsilon\lambda y(Fa)],x)$, it is not true that $\forall x([\varepsilon\lambda y(Fy)],x)$, so, a 's essentially being F ($[\varepsilon\lambda y(Fy)],a$) will not be generalised away, even though a 's 'essentially being such that a is F ' is generalised away. This is compatible with both the property-based and proposition-based characterisations of essence presented by Fine. The property corresponding to $[\varepsilon\lambda y(Fa)]$ is generalised away from a 's essence, but the property corresponding to $[\varepsilon\lambda y(Fy)]$ is not. This means that a is still essentially F . Likewise, the proposition attributing the former predicate to every object is true, whereas the proposition attributing the latter predicate to every object is false. As such only the former, and not the latter is generalised away from a 's essence.

Fine anticipates a response to Case 3 along the lines of generalisation. He claims that whilst excluding essential properties that are held by everything would resolve the problem, it could be easily reinstated by "conjoining the given degenerate essential property with one which was not degenerate."¹⁶² This is true. Instead of being essentially such that $2+2=4$, Socrates is essentially such that he is human and $2+2=4$ (precisely the kind of property we confronted Fine with in Section Three).

A potential defence against this variation of the problem comes from Della Rocca.¹⁶³ Della Rocca's characterisation of essentialism includes a generalisation approach that precludes any trivial property from featuring in the essence of an object. He introduces a notion of triviality which, whilst not explicitly writing with Fine's work in mind, could be taken to circumvent this revival of the problem.¹⁶⁴ According to Della Rocca, an object has a property trivially if it is either universally instantiated, or follows directly from the object having some other property that is universally instantiated. This means that, for example, an object a will not be

¹⁶¹ Whilst Fine's predication form is intended for the expression of essentialist claims as understood by Fine, there is an appropriate understanding of these sentences in purely modal terms in the form of Wiggins' *de re* must (1976). Here, the 'essentially' operator is interpreted as a necessity operator ranging over the predicate (in line with how Wiggins thinks *de re* necessity should be understood).

¹⁶² Fine (1994), p.6

¹⁶³ Della Rocca (1996)

¹⁶⁴ Della Rocca (1996), p.3, Wildman (forthcoming) presents and dismisses a response along these lines, attributing it to Della Rocca.

essentially identical to *a*, because this follows directly from the property of self-identity, which is universally instantiated. However, as Wildman points out, this fails to provide sufficient response to the revived problem; such a rule would result in the exclusion of all candidate essential properties.¹⁶⁵ For example, consider Socrates' being human. This follows directly from the universally instantiated property of being human if Socrates. Everything is such that it is human if it is Socrates, and because it follows from a universally instantiated property it must be excluded. The same can be reconstructed for any candidate property. This is, suffice to say, undesirable.

What of Fine's own generalisation response? Fine claims that the inability to rigorously demarcate constitutive essence is not a problem because consequential essence can be rendered adequate once the universally instantiated properties have been generalised away. Fine's generalisation proposal does not go as far as Della Rocca's, and so does not inherit those problems, but it is in no better position than the generalisation responses he anticipates in *Essence and Modality*. After all, being essentially such that he is human and $2+2=4$ follows from Socrates' essence under the neo-Aristotelian interpretation, just as it does from the modal one.

The neo-Aristotelian essentialist can avoid this jeopardy by adopting the characterisation of logical concepts as inferential in nature. Further, they can adopt my proposed definition of constitutive essence, and not worry about such problems at all. It is not clear what relief might be gained for the modal essentialist by adopting Fine's characterisation of logical concepts, but perhaps they might avoid the problem by making their own distinction between constitutive and consequential essence. First off, it is clear that the modal essentialist cannot adopt my proposed definition of constitutive essence because it incorporates the distinctly neo-Aristotelian TIVON relation. What of the original form that Fine presents? According to this:

¹⁶⁵ Wildman (forthcoming), p.3

It is **constitutive** of *a*'s essence that it is F iff *a* is essentially F, and that *a* is F is not a logical consequence of *a* being essentially φ (where φ is any 'more basic' property *a* essentially has).¹⁶⁶

Adopting the notion of constitutive essence as presented here would successfully exclude the problematic propositions. However, it inherits the problems that originally led Fine to favour the consequential account over the constitutive. It is not clear how to distinguish between the 'more basic' properties in an object and the others. This problem seems even more pronounced for the modal essentialist, because all claims necessitated seem just as basic as each other.

Furthermore, the modal essentialist taking this form of response may be accused of presupposing the pre-theoretical content of the notion of essence that Fine presents. By distinguishing between the properties that tell us what the object is, and those that merely follow from the 'proper' essential properties, the modal essentialist risks assuming a level of fine-grainedness that necessity is not fit to bestow. By distinguishing between the constitutive and consequential essence in the modal account, the essentialist is making essence more fine-grained than any intensional notion. Does this not mean that essence is no longer intensional, but hyper-intensional? If an adequate selection criterion can be found that can distinguish between the constitutive and consequential essence and that can be expressed in modal terms then this does not entail a move to hyper-intensionality (and neo-Aristotelian essence), but rather a restricted intensional notion (much as the conditional modal account provides). Such a criterion can be found. Here I suggest that the modal essentialist can adopt the constitutive notion of essence, defined thus:

CONSTITUTIVE*: Proposition P is part of *a*'s constitutive essence iff P is true at every world at which *a* exists, and at every world where *a* exists, *a* and *a* alone is the thing that makes P true.¹⁶⁷

¹⁶⁶ Fine (1995a), p.57

¹⁶⁷ Depending on one's metaphysics of properties, this may be problematic. If to have a property is to instantiate a universal, then there will never be one thing in virtue of which that thing has the property (except for homological properties, such as the universal of being a universal).

It makes sense to think that in a world w , some propositions will be true in virtue of some things, rather than others. For example, if it is true at w that the Eiffel Tower exists, it is true because the Eiffel Tower (or its counterpart in w) exists in w . Likewise, if it is true that there are infinitely many prime numbers, it is true because of the prime numbers.¹⁶⁸ Note that the ‘truth in virtue of’ relation here is different from the TIVON relation that Fine uses. The relation I intend is the kind of relation by which Socrates is what makes it true that “Socrates is human”, but at the same time is not sensitive in the way that Fine’s relation is, so Socrates is also what makes it true that “Socrates is bearded”. Socrates makes the former true in every world that he exists in, but does not make the latter true in every world he exists in. However, in the worlds in which it is true, it is true because of him.

Singling out the propositions that are true because of Socrates in every world he exists in singles out those propositions that make up his constitutive essence. What’s more, those propositions are selected for the right reasons rather than by some arbitrary or ad hoc process. Far from smuggling hyper-intensionality into the account, such an approach is in the spirit of modal essentialism. I take the motivation for modal essentialism to be the desire to account for certain counterfactuals we take to be true about objects, specifically those about how that object could or could not have been. The essence of an object can be understood as its modal profile. However, what modal essentialist accounts currently identify is the modal profile of the worlds singled out by an object’s presence in them. When I ask what change Socrates could survive, I am asking specifically about Socrates (and perhaps his counterparts). I am not asking about the worlds Socrates exists in. By concerning ourselves with just the propositions that are true because of Socrates in each world he exists in, we identify *his* modal profile, not that of the world. In addition, adoption of this approach would allow one to avoid the potential counterexample to modal essentialism whereby ‘Socrates is a man’ is true, but for reasons other than Socrates’ being there (e.g. worlds that contain a contradiction from which anything can be inferred).

¹⁶⁸ Or whatever nominalist explanation one favours.

By adopting an appropriate definition of constitutive essence such as the one presented above, the modal essentialist can respond to Case 3 by excluding the problem cases because they do not feature in the constitutive essence of the thing in question.

4.2 Case 2

Case 2 claims that because identity and non-identity claims are necessarily true, Socrates is essentially distinct from the Eifel Tower according to the modal essentialist. Fine denies that the essence of Socrates should feature the Eifel tower, and so modal essentialism is extensionally inaccurate. Whilst properties like ‘essentially being such that $a \neq b$ ’ ($[\epsilon \lambda x(a \neq b)], a$) can be generalised away (because $\forall x([\epsilon \lambda y(a \neq b)], x)$ is true), the same cannot be said for a ‘essentially being distinct from b ’ ($[\epsilon \lambda x(x \neq b)], a$). This remains an (according to Fine at least) undesirable essential property of a . It cannot be generalised away because not everything is such that it is distinct from b , namely, b is not distinct from b .

However, if the constitutive/consequential distinction is defensible, then it might be the case that being distinct from the Eifel Tower could be construed as a part of Socrates’ consequential essence, and could thus be safely ignored if we restrict our attention to his constitutive essence. If this is the case, then the modal account is left, like the neo-Aristotelian, depending on the defensibility of constitutive essence (with the further requirement that it is an acceptable move for the modal essentialist). If the proposed definition of constitutive modal essence given above is defensible (and can suitably place such claims in the consequential essence of an object) then these criteria are satisfied. However, first I show that the distinction as defined by Fine is of no avail to the modal essentialist.

Non-identity statements are necessary if true. Such claims are therefore essential to objects according to modal essentialism. This means that to avoid the problem using the constitutive/consequential distinction, we must admit that it is essential to a that $\langle a \neq b \rangle$, but contest that $\langle a \neq b \rangle$ is merely consequentially essential to a . Under

Fine's definition of consequential essence, $\langle a \neq b \rangle$ is part of a 's consequential essence iff it is a logical consequence of some more basic part of a 's essence.

Any particular instance of a being essentially distinct from something (excluding cases such as the distinctness of points in Euclidean space, where it is arguably constitutive) follows from a being essentially distinct from everything that is not a ($\forall x(x \neq a \rightarrow ([\exists \lambda y(y \neq x)], a))$). Furthermore, it follows from the self-identity of a that it should be distinct from everything else, thus placing the claim that a is distinct from b in the consequential essence of a if it is part of the constitutive essence of a that it is self-identical.¹⁶⁹

There is a problem with using Fine's definition in this approach. In order to avoid the self-related property regress, Fine suggests that we take the reflexive property as a more basic part of an object's essence, as it is the basis by which the self-relating property is included. However, in accordance with the generalisation response Fine gives in order to render consequential essence suitable for his purposes, any property that is universally instantiated will be generalised out from the essence of any object. Self-identity is a universally instantiated reflexive property, and so is generalised away, but without being essentially self-identical, a cannot be essentially such that it is identical to a ($[(\exists \lambda x)(x = a)], a$). As such, a can neither be essentially self-identical, nor essentially identical to a .

Now consider my proposed constitutive/consequential distinction. The necessity of non-identity statements ensures that $\langle a \neq b \rangle$ is still essential to a . As such, it must be demonstrated that being essentially distinct from b is only part of the consequential essence of a . To do this, it must be shown that $\langle a \neq b \rangle$ is true in every world in which a exists, but is not true solely because of a in all of those worlds.

It is worth noting that, unlike the original distinction, in the current distinction the consequential essence of an object does not depend on the constitutive essence of that object. According to the definition of constitutive essence under consideration, that a is self-identical is a part of a 's constitutive essence because $\langle a = a \rangle$ is true

¹⁶⁹ Not everyone would agree that self-identity is essential to a . The existence of such disagreement lends weight to my interpretation of the debate.

because of a in every world in which a exists. If the modal essentialist chooses to generalise away universally instantiated properties from objects' essences, then it is not essential to a that it is self-identical. However, this has no effect on whether $\langle a \neq b \rangle$ is part of a 's consequential essence. If a is necessarily such that $\langle a \neq b \rangle$, but this is not solely because of a in every world in which a exists, then $\langle a \neq b \rangle$ is part of a 's consequential essence regardless of what is in its constitutive essence. It is not clear why $\langle a \neq b \rangle$ would be true because of a alone. If it is not, then it is part of the consequential essence of a , and so can be ignored. If it is true because of a alone in every world containing a then it is part of the constitutive essence of a , and Case 3 stands.

If it is constitutively essential to a that it is distinct from b , what harm does this do? Fine's arguments rest on the assumption that if the results of modal essentialism differ from the pre-theoretical understanding of essence that he has in mind then they are shown to be extensionally inaccurate, and so false. In discussing Case 4, I consider the claim that this is not as problematic as Fine takes it to be.

4.3 Case 4

Case 4 pointed out that under the conditional modal account every object exists essentially because it exists in every world in which it exists. Fine claims that this is absurd, and as such modal essentialism must be mistaken. How we should view this criticism depends on how we interpret the larger debate between modal and neo-Aristotelian essentialism. There are two ways in which we might interpret this debate. First, we might see the two positions as trying to provide an explanation of a common notion of essence, one of which we already have an understanding. In such a dispute there is a base of facts, and an important desideratum for a theory is to be able to match those facts with what the theory predicts. An accusation of extensional inaccuracy, where the facts the theory predicts differ from the pre-theoretical base, is a serious one. Fine sees the debate in this way. He takes the base of facts to be supplied by a pre-theoretical understanding of essence, and claims that modal

essentialism cannot match them. Second, we might see the two positions as presenting rival notions that account for our limited pre-philosophical opinions and natural language usage, but do not assume a well understood pre-theoretical notion. Interpreted as such, an accusation of extensional inaccuracy does little more than highlight the difference between the notions the theories present.

Fine accuses the conditional modal essentialist of predicating objects with essential existence, and it is true that they do this. Fine takes this to be problematic because he has a pre-theoretic idea of what such a claim amounts to. However, accusing the modal essentialist of making a predication that is only problematic by the neo-Aristotelian essentialist's light seems to be little more than begging the question against the modal essentialist. Within the context of the second interpretation of the debate, essential existence must be judged by the modal essentialist's lights. For the modal essentialist, ubiquitous essential existence is exactly what we should expect. It is simply the claim that every object exists in every world in which it exists. It is a trivial consequence of the theory, and is neither of any particular interest, nor obviously problematic. Under neo-Aristotelian essentialism, essential existence is problematic, but we are not committed to it. Under modal essentialism, we are committed to essential existence, but it is not problematic.

However, this is not the project that Fine is engaged upon. Fine has a pre-conceived understanding of what it is for something to essentially exist (which he considers to be informed by our pre-philosophical opinions), and he thinks this pre-theoretical understanding of essence renders the essential existence of everyday objects absurd.¹⁷⁰ As such, any account of essence that says that everyday objects exist essentially is not only extensionally inaccurate, but committing an absurdity.

What makes the essential existence of everyday objects absurd? Fine does not make it explicitly clear why it would be problematic for Socrates to essentially exist. Presumably he thinks that it conflicts with our pre-theoretical understanding of

¹⁷⁰ Fine (1995a), p.58

essence that he is trying to present, and that it is thus obvious. He merely asserts that “we do not want to say that [Socrates] essentially exists.”¹⁷¹

Here are some potential worries that may motivate Fine’s repudiation of essential existence:

1 – Including existence in the essence of a thing tells us nothing about *what it is to be* that thing, at least, it tells us nothing unique about what it is to be that thing. This is true, but existence differs from set membership and non-identity claims. Existence is at least relevant to a thing in a way that the others are not. It is certainly true that Socrates could not be Socrates without existing.

2 – One may claim, as Kant does, that existence is not a real property, but rather a precondition for having properties, and as such is not an appropriate constituent for an object’s essence.¹⁷² It is unlikely that this is what Fine is thinking, as he seems to be willing to countenance some things having essential existence (God, for example). Furthermore, if existence is not a true property, then it need not be included in these considerations, and the modal essentialist would not be able to predicate it of everyday objects, or indeed at all.

3 – Fine may be worried about some kind of Anselmian ontological argument that would allow the actual or necessary existence of a thing to follow from its essentially existing.¹⁷³ This seems like an unlikely interpretation of Fine’s motivation, for the same reasons that modern philosophers of religion are generally not worried about ontological arguments. Furthermore, whilst *a* being essentially F entails that it is necessarily F, the intended form of necessity is only the weakened Kripke necessity, whereby *a* is necessarily F if it exists. There is no sense in which accepting the neo-Aristotelian essential existence of *a* commits one to the actual existence of *a*.

¹⁷¹ Fine (1994), p.5

¹⁷² Kant (1781)

¹⁷³ Anselm (1078)

4 – One may be concerned that if everything essentially exists then there is no distinction between that which essentially exists and that which exists only accidentally. This concern rests on equivocation about the meaning of essentiality. If ‘essentially’ were synonymous with ‘necessarily’, then indeed there would be a problem with a position that could not distinguish between those things that exist essentially and those that do not. However, this is not the neo-Aristotelian position, and I can see nothing built into our understanding of essence (as presented by Fine) that supports the need for such a distinction, let alone any cause to see the lack of one as absurd.

Fine faces a dilemma. Either there is no one well understood notion of essence and the different accounts are presenting rival notions that might explain what we mean when we make essentialist claims, or both accounts are trying to provide an account of a well understood notion. If one chooses the former then Fine is begging the question against the modal essentialist because whilst it is true that everything exists essentially, this is not a problematic, or even interesting conclusion considering what this means according to the modal essentialist’s position. If one chooses the latter then it is not even clear what is supposed to be problematic about essential existence. There is no obvious source or explanation of the absurdity that Fine ascribes. Of course, the provision of a good reason why essential existence is problematic would reinstate the criticism’s force. However, it is not clear what form such a reason would take.

Prima facie there is no evident data to support the claim that we possess a pre-theoretical understanding of essence sophisticated enough that it informs us on whether or not objects essentially exist. Views on this appear to be informed solely by theory, in which case essential existence is a weak tool by which to criticise the modal essentialist. The same can be said for the alleged problem from Case 2. If it is the case that objects are essentially distinct from each other, this is only really problematic under the first interpretation of the debate. A significant part of why Fine finds it problematic is because of his views on dependence. According to Fine,

an object is ontologically dependent on those objects that feature in its essence. This makes it problematic, as it would be odd for Socrates to be ontologically dependent on the Eiffel Tower. However, such an understanding of dependence is contingent upon accepting Fine's essentialism, so again it seems that the problem begs the question in Fine's favour.

4.4 Case 1

Case 1 claimed that it is essential to {Socrates} that it contain Socrates, but not essential to Socrates that he be a member of {Socrates}. Since the modal essentialist cannot account for this asymmetry it is extensionally inaccurate. The first concern with this kind of example is that it requires the use of contentious and unusual metaphysical entities such as sets. One might think that if the concerns highlighted are legitimate then they should be demonstrable through more everyday examples.

Fine claims that this is not problematic because he does demonstrate his point through everyday examples. Fine's other cases involve objects like Socrates and the Eiffel Tower, both of which are thoroughly mundane. The case of Socrates and {Socrates} merely demonstrates the desired asymmetry well, and for the reader who is not convinced by the subject matter, Fine has three other cases that should firmly press the point to satisfaction. However, the above considerations demonstrate that Fine's other cases are far from conclusive, and left on its own, the Socrates/{Socrates} case loses much (though not all) of its force.

Fine's approach is to claim that the intensional notion of essence is not co-extensional with our pre-theoretical notion of essence. However, if the modal essentialist's understanding of essence is as the modal profile of an object, then set membership doesn't seem like an unreasonable topic at which to bite the bullet and reject the pre-philosophical opinions that Fine considers paramount.

This links back to the metaphysical contentiousness of the example at hand. Not only does the criticism require that one believe in such things, but it requires that our pre-theoretical understanding should be sufficient to inform us about the essences of

such things, or indeed how the essences of everyday objects relate to them. However, it seems outlandish to think that this should be the case. After all, there is little reason to think that intuition should be anything more than a guide when considering the metaphysics of sets and personal identity. Intuition furnishes us with no decisive information about set theory or advanced metaphysics, just as it furnishes us with no decisive information about particle physics or anatomy.

5. Conclusion

Even if we interpret the debate on essentialism as between rival attempts to give metaphysical explanations for an accepted basis of pre-philosophical opinions, the cases Fine presents are not decisive in showing the inadequacy of the modal account. The modal essentialist can employ tactics (inspired by Fine's developments of his position) to respond to Cases 2 and 3. Case 4 is not evidently problematic, and Case 1 is based on assumptions for which pre-philosophical opinion cannot provide legitimate information.

How best to interpret Fine's work then? If we take Fine's four cases as definitive of a more fine-grained notion of essence, rather than descriptive of a pre-theoretical one, then the virtues of Fines' position come into sharp relief. Neo-Aristotelian essentialism integrates better with a wider metaphysical program. It has the potential to provide theories of necessity and ontological dependence. It also makes essence a more interesting topic for the metaphysician. Rather than being an uninteresting special case of necessity, essence plays a central explanatory role in our theory of the world.

Chapter Four: Antirealism about Neo-Aristotelian Essence

1. Introduction

In the previous chapter I presented neo-Aristotelian essentialism, as contrasted with modal essentialism. I suggested that whilst the Finean arguments against modal conventionalism are not decisive, there is still reason to favour the neo-Aristotelian conception in terms of theoretical utility. In considering antirealist accounts of essence this advantage increases. The neo-Aristotelian conception of essence provides a better explanation of why we think in terms of essence in the first place. This is of particular importance because for an antirealist account to work (in any area of discourse) it must be able to explain why we use that discourse (in this case, why we essentialise). That is, they must explain not only what it is to make essentialist claims, and why it is correct for us to do so when we make them successfully, but what we gain in our everyday lives by making first-order object language essentialist claims. Neo-Aristotelian essentialism provides a candidate explanation for why we essentialise in that the definitional character of essence plays an explanatory role in how we understand the world we live in. It does so in a way that modal essentialism does not.

Chapter Two established the viability of an antirealist account of modal essence via Sidelle's neo-conventionalism; in this chapter I explore the potential for an antirealist account of neo-Aristotelian essence using the Sidellean framework. I conclude that such an account is both available, and a viable option for the antirealist about essence.

In Section Two I give a recap of the neo-Aristotelian essentialist position. In Section Three I present the different options available when pursuing an antirealist account of neo-Aristotelian essence. In Section Four I explore how a broadly Sidellean mechanics can be employed to construct a theory for each of the options

presented in Section Three. In Section Five I propose an antirealist theory of neo-Aristotelian essence, and dispense with the need to discuss essentialism using propositions and the mysterious TIVON relation. In doing so I address the concern that the account presented in this chapter is not in fact one of neo-Aristotelian essence, but rather of modal essence in disguise. In Section Six I motivate the adoption of the antirealist accounts presented in this chapter and in Chapter Two over their realist counterparts. The advantage these positions have becomes evident when we pay especially close attention to the essences of kinds, something that up until this point has largely been secondary to the essences of individuals. Using Correia's discussion of objectual and generic essence, I argue that the essences of kinds are best thought of in terms of generic essentialist claims, and that the antirealist positions presented here are better suited to accounting for generic essentialist claims.¹⁷⁴

2. Neo-Aristotelian Essentialism

In the previous chapter, we settled on a formulation of neo-Aristotelian essence of the form:

ESSENCE: *a* is essentially *F* iff the proposition $\langle Fa \rangle$ bears TIVON to *a*, and *a* alone.

where TIVON (True In Virtue of the Nature Of) is a unanalysed relation between propositions and objects. The essence of an object is the collection of propositions that bear TIVON to that object.

TIVON is the key primitive of the theory; essentialist facts are grounded entirely in instances of TIVON holding between propositions and objects. Whilst TIVON was taken to be unanalysed, we were able to present an intuitive understanding of the relation so that we might have some idea of when it holds and what essentialist facts it presents us with. A proposition bears TIVON to an object (or group of objects) when that proposition is true because those objects are the objects that they are. The

¹⁷⁴ Correia (2006)

propositions that bear TIVON to an object form a subset of the propositions that are made true by that object. Whilst <Socrates had a beard> may be true because Socrates existed and had a beard, that proposition does not bear TIVON to Socrates. However, <Socrates was a human> does bear TIVON to Socrates because there is something about Socrates being Socrates (and not just existing) in particular that makes <Socrates was a human> true.

3. Options for Antirealism

In Chapter One we established three roles that convention might play in securing the truth of a sentence. These are conventional grounding types 1, 2, and 3. A sentence *P* is type 1 grounded in convention when *P* is true iff some convention *C* holds (where *C* does not feature in the meaning of *P* as in, for instance, “we hold convention *C*”). *P* is type 2 grounded in convention when *P* is true iff some other fact *F* holds, where *C* determines that the truth of *P* is contingent upon *F*. *P* is type 3 grounded in convention when *P* is true iff some fact *F* holds, where what counts as an instance of *F* is determined by *C*. Having presented neo-Aristotelian essentialism we are able to consider what form an antirealist account of essence might take. There are several potential forms available. Here I present those options that involve accounting for essence using type 2 or type 3 conventional grounding. For this discussion I use the term ‘robust’ to indicate that something is a feature of convention-independent reality, and not in any way grounded in the way we think about or interact with the world.

Deciding on which antirealist option to take is deciding what kind of conventional grounding is occurring, and what it is that is being grounded. Doing so does not constitute a complete theory, just as the neo-conventionalist positions examined in Chapter One do not. In order for the theory to be complete, we must take this first decision, and then introduce a mechanism by which the conventional community allegedly determines the subject of, and then performs, the conventional grounding, and that explains why we essentialise in the first place. This is the task of

Section Four. For the purposes of this section, as in Chapter Three, I use the propositional conception of essence presented by Fine, explaining essence using the TIVON relation. Here I examine the different forms that antirealism about neo-Aristotelian essence might take.

There are three points in the neo-Aristotelian account of essence that are potential areas of variance. One might think that any of: the TIVON relation, the things themselves, or the reduction that makes essence dependent on them, might fail to be robust. There are eight possible combinations of these three factors being conventional or robust together. These combinations are presented as options 1 to 8 in the table below.

Option	TIVON	Objects	Reduction
1	No TIVON	No Objects	No Reduction
2	No TIVON	No Objects	Robust Reduction
3	No TIVON	Robust Objects	No Reduction
4	No TIVON	Robust Objects	Robust Reduction
5	Robust TIVON	No Objects	No Reduction
6	Robust TIVON	No Objects	Robust Reduction
7	Robust TIVON	Robust Objects	No Reduction
8	Robust TIVON	Robust Objects	Robust Reduction

Not all of these options are worthy of our attention. Option 8 corresponds to realist neo-Aristotelian essentialism as explored in Chapter Three, so it can be excluded. Furthermore, in Chapter One I concluded that type 3 grounding in convention commits us to type 2 grounding in convention.¹⁷⁵ As such I shall exclude options 2, 4, and 6 from consideration because they involve type 3 grounding, but not type 2. This leaves us with options 1, 3, 5, and 7. For the purposes of discussion I will name the remaining options by their defining features. Because all of these options feature no-reduction, I will remove this feature from the names of all options bar option 7,

¹⁷⁵ Admittedly not in all cases (see Chapter One for potential counterexamples), but in the present case I take this rule to hold.

which I will simply call the no-reduction option. The remaining options will be named like so: option 1 is the no-TIVON no-objects option; option 3 is the no-TIVON robust objects option; option 5 is the robust-TIVON no objects option. I discuss each option separately.

3.1 No-Reduction

According to the no-reduction option *a* is essentially F iff the proposition $\langle Fa \rangle$ bears TIVON to *a*. Both *a* and TIVON are robust, but the biconditional is conventional (that is, the dependence of *a* being essentially F upon $\langle Fa \rangle$ bearing TIVON to *a* is a matter of convention). Essence is type 2 grounded in convention because whilst the explanans is robust, the explanandum's relation to it is determined by convention. This is similar to the way one might think a painting is considered to be a work of art. One might hold that painting *a* is a work of art iff *a* has feature F.¹⁷⁶ Painting *a* is robust, as is *a*'s possession of feature F. It is only *a*'s being a work of art that is not robust, and it is conventional that it is *a*'s being F (as opposed to its being G, or H, etc.) that makes it a work of art. As such, *a*'s being a work of art is type 2 grounded in convention; the same is the case for essence under the current option.

Whilst a coherent position may be formed out of this option, it would forfeit the original motivation for a (particularly an antirealist) neo-Aristotelian account of essence. Saying that *a* is essentially F iff certain robust conditions hold is unappealing for the same reasons that modal essentialism is. Modal essentialism, whilst not defeated by Fine's criticisms, fails to bestow any particular metaphysical purpose on essence. In the no-reduction option, the conventions that ground the reduction for neo-Aristotelian essence seem to do even less, merely giving a meaning to the terms by which we perform essentialist predication. Our conventions do little more than determine when we apply the term 'essentially'. As such, there is no evident reason why we should essentialise at all. It is not obvious what advantage is bestowed upon a conventional community that adopts such conventions.

¹⁷⁶ Obviously this is a vastly simplified analysis, and should only be taken as a toy example.

Furthermore, if both the objects and the TIVON relation are robust, then there appears to be robust metaphysics going on that is doing the explanatory work required of essence. In which case it's not clear why we should hold conventions to do the same work. Essence appears to be either a realist notion (and our conventions are merely determining the meanings of terms), or conventional, but under-motivated. This is far from the results being sought by the project. As such, I will not consider this option any further.

3.2 Robust-TIVON No-Objects

According to the robust-TIVON no-objects option a is essentially F iff the proposition $\langle Fa \rangle$ bears TIVON to a , but the object a is in some sense conventional. A consequence of this option is that (*ex hypothesi*) one will be committed to antirealism about objects. There are two ways of interpreting the conventional nature of objects. On the one hand, the world may contain no objects, and we conventionally carve some out of the pre-objectual stuff in much the same way that Sidelle suggests. Under this interpretation, the objects are themselves conventional. On the other hand, there might be a great many objects (perhaps in the same way that would result from universal composition being true, whereby there would in fact be many more objects than those that we recognise or care about) and we conventionally select a subset of the existent objects that are of interest to us, and we populate our ontology with them. Under this interpretation, the objects themselves are robust, but their relevance and interest to us is conventional.

If we take the former option, whereby we carve objects out of the pre-objectual stuff of the world, then it is hard to see how TIVON can be robust, seeing as it is a relation between propositions and objects, and only the relata on one side (the propositions) exist in a robust sense. One might hold a principle to the effect that a relation can only be as robust as its least robust relata. This seems plausible, and if accepted would commit us to TIVON not being robust. This is not necessarily a bad

position (indeed, as I argue later it makes for a compelling one), but it is not the position under consideration, so I shall postpone discussion of it until later.

If we opt for the latter option, whereby there is a plenitude of objects from which we select the ones of interest to us, then it is hard to see how essence is truly conventional. It seems more like essence is robust for a realm of real objects that is far more populous than we actually think, and we merely use convention to pare down the size of that realm to something more manageable. It is our selection of which objects are of interest that is conventional, not the essences of those objects. One might find such a position appealing when compared to a more full-blooded antirealism, but to explore such an account fully would be orthogonal to the goals of this project, so I shall set this option aside. As such, we can conclude that (for the purposes of this project) in order for a no-objects option to be considered properly, it must be combined with antirealism about TIVON.

For these reasons I will abandon the robust-TIVON no-objects option and the no-Reduction option. In the next section I narrow my focus to the two remaining options. These are the no-TIVON robust-objects option, and no-TIVON no-objects option. I present those options here.

3.3 No-TIVON Robust-Objects

According to the no-TIVON robust-objects option a is essentially F iff the proposition $\langle Fa \rangle$ bears TIVON to a , and a is a robust object, but where convention determines what propositions bear TIVON to a . Under this option, TIVON is a function from the set of propositions that are made true by an object to a subset of that set that are taken to be definitive of that object. Essence is type 3 grounded in the conventional TIVON relation because whilst a being essentially F depends on $\langle Fa \rangle$ bearing TIVON to a , whether or not that relation holds is a matter of convention.

A consequence of this option is that the essentialist is no longer allowed to take TIVON as primitive. In one sense this is obvious. If TIVON is conventional then of

course it cannot be metaphysically primitive; it is not even metaphysically robust. However, there is a further sense in which TIVON cannot be primitive; let us call this sense ‘conceptually primitive’. For something to be conceptually primitive is for there to be no explanation of how it works; in effect it would be brute in its application. The closest thing to a conceptually primitive conventional notion might be an analytic one. As this project is built on the assumption that type 1 grounding in convention (the kind of conventional grounding involved in analyticity) has been discredited, the conventionalist is not in a position to take TIVON as conceptually primitive. This means that for the antirealist, the workings of TIVON cannot be a mystery with no explanation (as they can be for the realist). TIVON cannot be unanalysable, but rather must be subject to a rule that we hold by convention that governs its application. As such, the conventional mechanism developed for this option must be able to provide an account of how TIVON works, and when it holds between potential relata. This requires a rule by which the relata of a particular instance of TIVON can be determined, and a convention we hold that secures the application of this rule.

3.4 No-TIVON No-Objects

According to the no-TIVON no-objects option a is essentially F iff the proposition $\langle Fa \rangle$ bears TIVON to a , where there is no robust object a , and convention determines what propositions bear TIVON to a . This account bears the strongest similarity to Sidelle’s conventionalism in that it has the same broad-reaching antirealism about objects. It has the same consequences for TIVON as the no-TIVON robust-objects option, but with the addition of antirealism about objects.

Having narrowed down the options for an antirealist essentialism to two (the no-TIVON robust-objects option and the no-TIVON no-objects option), I proceed to consider how a conventional mechanism might be introduced to the account under either option.

4. Conventional Mechanism

The conventional mechanism used in this chapter to present an antirealist theory of essence is inspired by that used by Sidelle. The final version of the Sidellean mechanism (as developed in Chapter Two) took the following form. As per modal essentialism, an object is essentially *F* iff it is necessarily *F* if it exists. The modified Sidellean account of necessity takes the form:

NECESSITY: A proposition $\langle \text{nec}(P) \rangle$ is true iff $\langle \text{nec}(P) \rangle$ can be expressed by a sentence that is the conclusion of a valid inference from a sentence expressing the claim necessitated and an appropriate true GPI.

The GPIs mentioned in NECESSITY take the form:

GPI Schema: (x) (If x belongs to kind *K*, then (if p is x 's *P*-property, then it is necessary that x is p))

where what x 's *P*-property is is determined by what kind x belongs to. The GPIs are grounded in conventions of the following form:

GPI Conventional Schema: If ' x ' is a *K*-term then if p is the *P*-property of the thing denoted by ' x ', then ' x ' applies to something in any possible situation only if it is (has) p .

In this section I investigate how this mechanism can be adapted so that it can be implemented in the different options presented in Section Three. If the Sidellean mechanism can be used to provide an antirealist account of neo-Aristotelian essence comparable to that which it provides for modal essence, then (providing that whatever changes are required are not too substantial) the resulting theory of essence will be at least the equal of the modal essentialist position considered in Chapter Two. First I consider the no-TIVON no-objects option, and then the no-TIVON robust-objects option.

4.1 No-TIVON No-Objects

Presenting a Sidellean account of neo-Aristotelian essence that is antirealist about TIVON and objects is remarkably straight-forward. The GPI schema from Sidellean conventionalism can be adjusted so as to give a criterion that governs when TIVON holds between a proposition and an object, like so:

TIVON Schema: (x) (If x belongs to kind K , then (if p is x 's P -property, then the proposition that predicates p of x bears TIVON to x)).¹⁷⁷

When it comes to formulating the schema for the conventions we hold that ground TIVON Schema, the GPI conventional schema from Sidellean conventionalism is fit for purpose unchanged.

GPI Conventional Schema: If ' x ' is a K -term then if p is the P -property of the thing denoted by ' x ', then ' x ' applies to something in any possible situation only if it is (has) p .

By using this convention in our theory, we establish that we essentialise by identifying certain properties of objects and kinds, and then resolving not to consider anything else to be that object/kind in any situation unless it has those properties. Note that talk of objects and kinds here should not be taken as committed to realism, but as a less labour intensive way of talking of those parts of the pre-objectual stuff of the world to which we decide to apply object and kind terms. We can talk of an object a being F if we are talking about a situation in which we are willing to apply a , and where the stuff we apply a to is F . We establish what kinds and individuals there are, and what is essential to them, via rules controlling circumstances in which we are willing to apply the terms for them. For example if, in virtue of being human, one of Socrates' P -properties is his biological origins, then we resolve not to consider anything to be Socrates unless that thing is the son of Phaenarete and Sophroniscus. Likewise if, in virtue of being a chemical kind, one of Water's P -properties is its elemental composition, then we resolve not to consider anything to be water unless it

¹⁷⁷ This can be read as *just* the proposition $\langle a \text{ is } p \rangle$ so as to ensure that propositions like $\langle \text{Socrates is human and the Eiffel Tower is in France} \rangle$ do not bear TIVON to Socrates.

has the elemental composition H_2O . Note that whilst the conventions take the form of rules governing our application of terms for objects and kinds, what the conventions are really about is what an object/kind is and what we are willing to count as being that thing. That is, we are using conventions about language to provide us with the conventional essences of objects. We conventionally select the P-properties of objects and kinds, just as in Sidellean conventionalism, by selecting the key features by which we distinguish or single out the first instances of those objects/kinds. By selecting the P-properties that we do, we carve objects and kinds out of the pre-objectual stuff of the world.

This shows the extent to which Sidellean conventionalism bridges the gap between the modal and neo-Aristotelian interpretations of essence. According to Sidelle, the conventions that give us modality are the same conventions by which we divide the world up into objects and kinds. The conventions we use to carve the world also ground necessity. The modal distinctions are a by-product of our essentialising. Whilst Sidellean conventionalism was introduced in Chapter Two as an attempt to provide a conventionalist account of modal essentialism, the lack of additional work that is required to adapt his framework to suit neo-Aristotelian essentialism suggests how Sidelle's position (unintentionally) bridges the gap between the two positions. The reason for this is the strong link that Sidelle posits between modality and ontology. In making necessity so closely linked to the method by which we populate our ontology, Sidelle evokes the motivations that the original modal essentialists tried to satisfy: that the essence of a thing is synonymous with that thing's modal profile. However, in making essence central to the very existence of objects and kinds, the theory does justice to the definitional character of neo-Aristotelian essentialism. One might even think of the Sidellean account from Chapter Two as grounding necessity in conventional essences; the essence of an object or kind being the collection of its P-properties. In this light we can think of this chapter as focusing on the conventional essences themselves.

The no-TIVON no-objects option allows for an adequate account of essence, both providing the conventional grounds of essentialist claims and explaining how

and why we essentialise. However, just as Sidellean conventionalism about modal essence does, it commits one to antirealism about objects and kinds as well as about essentialist facts. One might think that this is to be expected. After all, the pre-theoretical understanding of essence, that things must have their essential properties if they are to be the things that they are, entails that things cannot exist without their essential properties. It makes sense that to be antirealist about essences would be to be antirealist about the objects and kinds that those essences define.

I think this reasoning is mistaken. It is one thing to be an antirealist about certain essences, for instance the essences of chairs, and for this antirealism to commit one to antirealism about such objects. It is quite another thing to be antirealist about the notion of essence altogether. On the first approach we accept the essentialist understanding of ontology, whereby for something to exist is for it to have an essence.¹⁷⁸ With this background assumption in place it makes sense to accept that antirealism about essences commits one to antirealism about objects. If I were to deny that anything but the natural kinds have essences then, under the above assumption, I would be committing myself to the claim that natural kinds are the only things that exist. However, a universal antirealism about essence, claiming that nothing at all has an essence, seems ill matched with such a characterisation of what it is to exist. If one claims that nothing has an essence then presumably the reasonable next step is to deny that there is such a thing as essence, and to excise the notion from our metaphysics (regardless of what things one thinks exist). If there is no such thing as essence in our metaphysics, then our metaphysics cannot make existence contingent upon essence.

The position being advocated in this project is not a universal application of the former strategy, whereby we have a realist understanding of essence but renounce each and every instance of it, but rather an adoption of the latter, whereby the realist notion of essence is renounced altogether. Rather than claiming that what it is for

¹⁷⁸ It is not clear that this is how all essentialists view ontology. Indeed, essentialism is usually given in the weak form of the claim that some objects have essences. However, this conception of ontology is clearly the one assumed in the reasoning that compels one to accept the inference from antirealism about essence to antirealism about objects.

something to exist is for it to have an essence, and then denying that there are any such essences, here we deny that there is anything more to essence than the way that we think about and act in the world. To claim that this commits us to antirealism about objects is to beg the question in favour of the essentialist conception of ontology, whereby essence furnishes reality with the joints at which we then carve. As such, whilst it is a legitimate dialectical move to adopt an antirealist position regarding objects and kinds, doing so solely because one is an antirealist about essence, and because there can be no kinds or objects without essence, is to make an assumption to which one is not entitled.

Of course, this is not to say that there must be an account whereby we can be antirealists about essence and realists about objects and kinds. It may be that no viable theory can fit into our conventionalist framework. The conclusion to draw, rather, is that the mere fact there are no robust essences is not sufficient grounds to infer that there are no robust objects. With this in mind, it is worth investigating the possibility of an account that is conventionalist about TIVON, but where objects and kinds are taken to be robust.¹⁷⁹

In what follows I investigate the viability of such an account. In doing so I am not rejecting or discarding the no-TIVON no-objects option. Indeed, I think that it turns out to be the better position. However, the goal of this chapter is not to conclude on which option is best, but rather to demonstrate that there are viable options available for the antirealist about neo-Aristotelian essence using a conventionalist framework developed from Sidelle's (or at least, something similar).

4.2 No-TIVON Robust-Objects

For a no-TIVON robust-objects approach to work, the account of essence must involve type 3 grounding in convention, whereby *a* is essentially *F* iff the proposition

¹⁷⁹ One potential, but ultimately unsatisfactory option would be to take an epistemological approach. Strictly speaking, this kind of antirealism need not commit one to antirealism about kinds and objects simpliciter, but only to those to which we have epistemic access. One could claim that there are kinds and objects, but that they are epistemically removed from us, and as such they are not the objects that we think exist.

$\langle Fa \rangle$ bears TIVON to a , but where what makes it the case that $\langle Fa \rangle$ bears TIVON to a is conventional. However, the application of TIVON will be governed by type 2 grounding in convention, whereby $\langle Fa \rangle$ bears TIVON to a iff some condition holds, where that condition features only robust elements, but where the biconditional itself is conventional. The task of this section is to present a principle that governs the conventional grounds for when TIVON holds, providing conditions for its application.

To present the principle that governs when a proposition bears TIVON to a , one might re-use the same principle as in the no-TIVON no-objects option.

TIVON Schema: (x) (If x belongs to kind K , then (if p is x 's P-property, then the proposition that predicates p of x bears TIVON to x))

There are two ways to interpret this schema. On the one hand we might take the P-properties of objects to be robust, and as such maintain that this schema represents the type 2 grounding of TIVON in convention (because the conditions in which TIVON holds are robust, but TIVON's being contingent on those conditions is not). This would mean that objects have the properties that we describe as P-properties, and those properties do have the special P-property status. The claim is that, despite the objects existing and having P-properties, it is a matter of convention when TIVON holds, and as such it is a matter of convention what the essential properties of an object are.

Alternatively, one might claim that whilst the objects themselves are real, the P-properties are not robust features of those objects. Whilst the objects would still have the properties themselves, their significance as P-properties would not be a robust feature of those objects. To take this option would mean that TIVON is also type 3 grounded in convention. This option claims that we conventionally select the P-properties of kinds and objects, and in doing so, we provide the type 3 grounds for TIVON's application (because TIVON's application is contingent on the P-properties, but what counts as a P-property is a matter of convention). In order for this approach to be viable for the no-TIVON robust objects option, we must show that identifying the P-properties as conventional does not commit one to antirealism about the kinds

and objects we associate them with. I first consider the type 2 option, and then the type 3 option.¹⁸⁰

4.3 Investigating the Type 2 No-TIVON Robust-Objects Option

This option claims that the P-properties (and their special status as P-properties) are robust features of objects and kinds (though it is less than clear what it means to be a P-property at this stage). The role of our conventions is to determine that it is the P-properties that are the properties of importance for essential predication, rather than any other subset of said object/kind's properties. In order to maintain this, our metaphysics must be able to say in what way the P-properties of an object or kind differ from those other properties of that object or kind. This is a difficult task. P-properties cannot be different from other properties just because they play a role in making an object what it is, or because the object has them because it is what it is. To make either claim would be to presume some form of realist essentialism. Alternatively, it cannot be the case that there is nothing special about them at all and that they are just an arbitrary grouping of an object's properties, because then there would be no reason why they have their special status as P-properties (after all, they don't do anything special; they are otherwise unremarkable). One might claim that they have their status simply in virtue of being the properties that we selected to be relevant to the application of TIVON. They are conventionally selected, and we could just have easily selected otherwise. This would provide an appropriate explanation, but as an answer it steps beyond the bounds of this option. The option under consideration is that of type 2 grounding; whether TIVON holds is contingent on the robust P-properties, but that dependence (and not what is depended on) is a matter of convention. To claim that whether TIVON holds is contingent on the P-properties, and that what counts as a P-property is a matter of convention is to enlist type 3 grounding in convention. The type 3 grounding option will be considered next.

¹⁸⁰ It is worth noting that the type 3 option does also involve type two grounding in convention. I call it the type 3 option because it is this that distinguishes it from the alternative.

If the P-properties are in some way closely linked to the object/kind being the thing that it is, and if this were to somehow avoid the presumption of realist essentialism, then the conventional aspect of the theory appears to be redundant. There are robust features of the world that are performing the theoretical role that essence otherwise would and, if epistemically accessible, explain why we essentialise. Under such a state of affairs it is hard to see how our essentialising would not 'link up' with the metaphysics of the P-properties, instead of conventionally grounding a whole other realm of discourse.

As such, it seems that the type 2 grounding option is not worthy of further pursuit. Next I consider the type 3 option.

4.4 Investigating the Type 3 No-TIVON Robust-Objects Option

The central claim of this approach is that the special status of objects' P-properties is conventional, but that this does not preclude those objects from being robust features of the world. It is not enough to merely claim that this is the case. It must be shown that the claim can be reconciled with the mechanics of the account. In particular, it must be reconciled with the conventions we hold that select which of an object/kind's properties are its P-properties. In Sidelleian conventionalism it was by selecting the P-properties that we carved out the kinds and objects. When we first individuated them we did so by selecting the properties by which we were to identify further instances of that kind. This option is not available in this instance. If the objects and kinds themselves are already robust features of the world, then the conventional selection of P-properties must perform some function other than dividing up and categorising the world. Remember that whatever instances of the P-properties objects and kinds have, they have them independently of us. Water has the elemental composition H_2O , and Socrates is the son of Phaenarete, regardless of whether we consider elemental composition or biological origin to be P-properties of chemical kinds and animals respectively.

Formulating a no-TIVON robust-objects, no-P-properties account requires that we answer four questions. Can (robust) objects exist without (robust) essential profiles? Can (robust) kinds exist without (robust) essential profiles? If the answers to one or both of these questions is ‘yes’ then what is the theoretical role of essence (that is, what does it contribute to our theory, be it by explanatory power, or other benefit)? If the answers to one or both of the first two questions is ‘yes’, then why do we essentialise? The answers to the last two questions may well be similar, but the questions are worth considering separately. If these questions receive satisfactory answers, then the task remains to produce a conventional mechanism whereby how we essentialise is explained, and how our behaviour grounds the essentialist facts the theory posits. Here I address each question in turn.

4.5 Can (Robust) Objects Exist Without (Robust) Essential Profiles?

In the above I argued that to assume that antirealism about essence commits us to antirealism about objects is fallacious. This argument is based on the claim that it is a significant metaphysical assumption to link the existence of an object to its having an essence. Making such a claim relies on an object’s identity conditions being (at least in part) constituted by its existence in certain counterfactual situations. This is because it is through determining what is the case in certain counterfactual situations that the essences of objects determine their identity criteria, and it is arguably through the identity criteria of objects that essence plays a role in their existence. Which counterfactual situations contain an object (or an appropriate counterpart of the object) indicates what kinds of modal change it can survive. If there are no essentialist facts (and so no counterfactual situations if modality reduces to essence in the way Fine claims), then only how an object actually is can be involved in an object existing as an individual distinct from the world around it.

If only the way the world actually is, and not the way it could or could not be, is relevant to the robust existence of objects, then one might think that a certain brute objecthood is available. If our theory of composition provides us with mereological

sums (for instance), then we might consider such sums to be objects. Brute objecthood would allow there to be objects even if there were no essences, so long as the work that would be done by essence is done by the brute objecthood of those objects. However, one might think that by providing a mereological criterion for objecthood all we are doing is selecting a kind of essential property. A second concern comes from the need to discern which object we are talking about in any circumstance. For instance, if I point to Socrates and say “this is the referent of ‘Socrates’” it is unclear whether I am referring to the aggregate of biological matter, or the person.¹⁸¹

It is important to remember that the only reason why such confusing scenarios arise is because of the application of Leibniz’ law to modal or essential properties. The reason why the person (call it ‘Socrates’), and the aggregate of biological matter that makes up Socrates (call it ‘Blobrates’) are considered to be separate objects is because they have a non-identical list of properties. This difference comes solely from their modal/essential profile. What distinguishes Socrates from Blobrates is differences in what kind of modal change they could survive. For example, Socrates could not have come into this world as a pile of sausages, whilst Blobrates could have. Likewise, Socrates (arguably) could have been made up of completely different biological matter, whereas Blobrates could not have. Without essential properties there can be no distinction between Socrates and Blobrates. By Leibniz’ law we should conclude that they are the same object. Whilst this results in a sparser ontology than one might otherwise have, this seems, in this case at least, advantageous.

4.6 Can (Robust) Kinds Exist Without (Robust) Essential Profiles?

It is not clear that there is any strategy equivalent to that given for objects that can be given for kinds. The only way by which the world can be divided into kinds is through the membership criteria of those kinds. The kind Cat is the grouping of

¹⁸¹ Ignoring even stranger potential objects such as Socrates’ skin, all of Socrates but his arms, the front half of Socrates, etc.

objects by certain properties. Whilst there may be many ways of interpreting how this works (whether by set membership, the instantiation of universals of some kind, etc.), I take it that these are differing mechanisms by which we do the same thing. Even if there were arbitrary kinds for every grouping of objects there would still be a membership criteria by which the kind is defined (even if this were an extremely gerrymandered criteria, like having the property of being Socrates, Marie Curie's chin, or Ganymede). It seems that any way of thinking about kinds relies on some form of essentialism. This is not parallel to our discussion of objects because the tasks we charge essentialism with are different for kinds than they are for objects. For an object, we are concerned with what counterfactual situations that object might have existed in, and perhaps what kinds of temporal change it can survive (for example, whether Socrates might have been born of different parents, or whether he could survive being teleported).¹⁸² For a kind, we are concerned with what it takes for something to be an instance of that kind (for example, whether this clear liquid before me is water).

This is not to say that counterfactual suppositions are irrelevant to the consideration of the essences of kinds. When I entertain the counterfactual situation in which the clear liquid before me is XYZ and not H₂O, I conclude that in such a situation the liquid is not water. However, such consideration does not inform me as to the essence of the kind Water in the same way that equivalent suppositions might inform me as to the essence of the individual Socrates. I learn about under what circumstances something is of that kind, not whether that kind would be the kind it is.

Note that there are some points of similarity between the essences of kinds and those of objects. Assuming that kinds are platonic abstracta, one might entertain the counterfactual situation in which the kind Water is concrete. Such a situation does not appear to really contain the kind Water, so we might conclude that (under the assumption that Water is a platonic abstract object) the kind Water is essentially

¹⁸² Whether one countenances the temporal examples in motivating essentialism is besides the point, here. Our essentialist beliefs do inform (to at least some extent) discussion of such questions.

abstract. As such, we can conclude that kinds have two types of essential properties, those relating to the members of the kind, and those relating to the kind as a thing itself.¹⁸³

Even if some criteria were found by which we could group objects into kinds, it is not obvious that this would be anything more than something akin to the use of P-properties in disguise. For example, one might appeal to objective similarity to group objects into kinds based on what objects they are most similar to. However, this tacitly relies on certain kinds of similarity being more important than others. For instance, grouping animal species by similarity relies on prioritising genetic makeup or perhaps ancestral lineage (over, say, geographical location), and grouping chemical compounds relies on prioritising elemental composition (over, say, phenomenal properties). As soon as the sharing of one property is prioritised over the sharing of others, that property is given status equivalent to that of P-properties.

As a last ditch attempt at finding a way to group objects into kinds, there is one approach that may satisfy our metaphysical needs, but not our epistemic needs. One might claim that objects are of the same kind because they have the same kind-property. What scorpions all have in common is that they all have the property of Scorpionhood. One might think that this avoids the concern of evoking P-properties by claiming that there are brute kind-facts (much as we claimed that there might be brute facts about what objects exist). Even if this solves the metaphysical problem in as far as it provides a criterion by which objects can be of the same kind without the use of P-properties, it does nothing to furnish us with a method by which we can tell what kind something is. There is no way for us to look at something and recognise the property of Scorpionhood without making an inference from other qualities we observe the object as having (e.g. having pincers and a stinging tail). However, making such an inference requires that there be some properties that are more important to the kind than others, thus defeating the point. In the absence of a suitable epistemology of kinds, this kind of response is far from satisfactory, and is little more than a philosophical trick. Whilst it may technically get the desired result,

¹⁸³ This point is of central importance, and discussed further, in Section Six.

it is hardly convincing as part of a larger theory. I proceed with the conclusion that, whilst antirealism about essence need not commit us to antirealism about objects, it does commit us to antirealism about kinds.

4.7 If the answers to one or both of these questions is ‘yes’ then what is the theoretical role of essence?

We are left in the interesting position that even if robust objects can be maintained despite antirealism about essence, the theoretical role that essence plays in Sidellean conventionalism is no longer appropriate. In Sidellean conventionalism the conventions that provide the essentialist facts are the same conventions that carve the world into kinds and objects. In the position being developed here half of that work has already been done for us by the world. However, whilst it may be the case that objects are already given to us by the world, it seems that kinds are still the workmanship of convention. As such, taking the theoretical role of essence to be in explaining kinds facts is a good place to start.

If the sole role of essentialist predication in our theory is in explaining kind facts, then one challenge that must be overcome is incorporating essentialist predication about objects, explaining what role this plays in explaining kinds. One potential avenue is that the essences of objects do not explain those of kinds, but rather are some kind of necessary consequence of them. The essential properties of objects are those that are important to them in virtue of what kinds they are a member of. One task for the proposed account in Section Five is to make essential predication of objects essential to the method by which we determine the essences of kinds.

4.8 If the answers to one or both of the first two questions is ‘yes’, then why do we essentialise?

If the theoretical role of essentialist facts is in explaining kind facts, then why we essentialise quickly follows. There are obvious advantages to thinking about the

world in terms of kinds. By grouping objects together in certain ways we can more easily ascertain whether something is dangerous, edible, useful, etc. One can imagine how vital such distinctions would be to early humans, and doing so by discerning what kind of thing something is, and then recognising that things of that kind are edible, is better than judging on a case by case basis.¹⁸⁴ Of course, the kinds Dangerous, Edible, and Useful are just as conventional as the other kinds. What is of importance is that we group things together in terms of our interest-relative priorities. As these priorities get more and more specific and specialised, so too do the kinds that we divide things up into. If there are no kinds in the world however, then we cannot make such distinctions without creating the kinds ourselves. We can explain our essentialising behaviour if we do it as a part of the way we categorise objects into kinds.

With rough answers given to the above questions we can proceed to develop an account of essence using the Sidellean mechanism that fulfils the no-TIVON robust-objects, no-P-properties criterion.

5. The Proposed Account

Something is essentially F iff it is F, and if F is a P-property of one of the kinds of which that thing is a member. This works in accordance with the TIVON schema presented previously.

TIVON Schema: (x) (If x belongs to kind K , then (if p is x 's P-property, then the proposition that predicates p of x bears TIVON to x))

Where x is a variable for objects or kinds, P-property stands for the kind of property by which we group the members of kind K together (e.g. elemental composition for

¹⁸⁴ Whilst this is obviously a simplification, it's worth noting that this story can be elaborated upon. For instance one need not be constrained by the pragmatic notions of this story. The distinctions that a conventional community make need not be set in stone, and can be subject to change. For instance, whilst it may be plausible to think that early people categorised the world in terms of things that were edible or dangerous, we are now more interested in what we might call 'deep explanatory features', like elemental composition, or particle charge. I see no reason why a similarly plausible story could not be given as to how we came to accommodate these further distinctions.

chemical kinds), and p is a possible value of the P-property (e.g. having the elemental composition H_2O).

The essential properties of objects are a result of the way that we divide the world up into kinds (for pragmatic purposes). The essential properties of kinds are the same. Not only do we divide objects up into kinds, but we also divide kinds up into kinds of kinds. Some kinds are such that they are less general instances of more general kinds. For example, the kind Human comes under the kind Mammal, which in turn comes under the kind Material Object. Where kind F is a less general instance of kind G , call F a *subkind* of G , and call G a *superkind* of F . I adopt the notation $F_{\text{sub}G}$ to show that F is a subkind of G , and $G^{\text{sup}F}$ to show that G is a superkind of F . A kind can have multiple subkinds and multiple superkinds. For example, Mammal has Human and Dog as subkinds, and Animal and Material Object as superkinds. Where F is a subkind of G and G is a subkind of H , F is a subkind of H ; where H is a superkind of G and G is a superkind of F , H is a superkind of F . As such, some chains of kinds will form nested hierarchies going from more general kinds at the top, down through the levels of subkinds to less general kinds at the bottom. A kind K is a subkind of another kind L when all of the things that fall under kind K fall under kind L .¹⁸⁵ Human is a subkind of Mammal, so Socrates is a mammal in virtue of being a Human.

One kind being a subkind of another should be distinguished from that kind *falling under* another. An object x falls under a kind K when it is of kind K . For example, Socrates falls under the kind Human, but Human does not fall under the kind Mammal (because the kind Human is a kind, not a mammal). Falling under is a relation that typically obtains between objects and kinds. Being a subkind-of is a relation that solely obtains between kinds and other kinds. However, kinds sometimes fall under kinds. The kind Human has various properties. Examples might include being non-empty, or being a species, etc. This means that it can fall under other kinds. Most notably, the kind Human is a kind, so it falls under the kind

¹⁸⁵ It is possible for kinds to be subkinds of each other. If $\text{nec}(\forall x(Fx \leftrightarrow Gx))$ then F is a subkind of G and G is a subkind of F . This may make the choice of terminology sound peculiar, but it is not problematic.

Kind. However, this does not mean that Human is a subkind of Kind. To say so would be to claim that everything that is a human is a kind. This is false.

So the kind Human falls under the kind Kind and it is a subkind of the kind Mammal. Kinds do not fall under the same kinds that they are subkinds of. There are no kinds that are mammals (though there are kinds that are subkinds of the kind Mammal), and there are no mammals that are kinds (though there are mammals that fall under certain kinds). Something *belongs to* a kind if it is an object that falls under that kind (e.g. Socrates belongs to the kind Human) or if it is a subkind and of that kind (e.g. the kind Human belongs to the kind Animal).

A heredity relation holds between subkinds and superkinds that extends to the objects that belong to them. A kind $F_{\text{sub}\{G_1 \dots G_n\}}$ inherits the P-properties of kinds $\{G_1 \dots G_n\}^{\text{supF}}$. The heredity relation is transitive, anti-symmetric, and reflexive. For example:

1. Socrates is a Human; being the son of Phaenarete and Sophroniscus is the biological origins of Socrates, and biological origins are the P-property of Humans.
2. Humans are Mammals; being warm blooded is a P-property of Mammals.¹⁸⁶
3. Mammals are Material Objects; being concrete is a P-property of Material Objects.

It follows from these claims that being warm blooded and being concrete are P-properties of Socrates, just as his biological origins are. As such, Socrates is essentially warm blooded, essentially concrete, and essentially the son of Phaenarete and Sophroniscus.

Having established the basics of the proposed position, certain clarifications are required with regard to how it applies to kinds.

¹⁸⁶ I use 'P-property' as a placeholder for whatever the reader prefers to think of as being the P-property. Whilst it becomes increasingly difficult to name the P-properties of more general kinds (succinctly and clearly, at least), this should not be taken as reason that there are no such properties.

5.1 Kinds

There is a potential for confusion when it comes to P-properties relating to the properties of kinds and those of individuals. Saying that Water is essentially H_2O , whilst true, seems to be a different claim from saying that a specific quantity of water is essentially H_2O . How the GPI mechanism allows us to secure the essences of objects and kinds requires scrutiny regarding how the transition from the essences of kinds to the essences of individuals is made. As an example we can present a hierarchy of GPIs.

First, a point of clarification is required. The kind Water is a subkind of the kind Chemical (because everything that falls under Water falls under Chemical), but does not fall under the kind Chemical (because the kind Water is a kind, not a chemical). However, the kind Water *does* fall under the kind Chemical-Kind (because Chemical-Kind is the kind of kind that the kind Water is), and is not a subkind of the kind Chemical-Kind (because not everything that falls under Water falls under Chemical-Kind (e.g. the water in my glass is not a chemical kind, it is not any kind of kind)). As such, I will proceed with the hierarchy of Water, Chemical, and Material.¹⁸⁷

If Water is a subkind of Chemical, which in turn is a subkind of Material, then according to the Sidelleian mechanism the first two links in the resulting chain might look like:

MATERIAL: If Chemical belongs to Material, then (if Elemental Composition is Chemical's Primary Explanatory Feature, then the proposition that predicates Elemental Composition of Chemical bears TIVON to Chemical).

CHEMICAL: If Water belongs to Chemical, then (if H_2O is Water's Elemental Composition, then the proposition that predicates H_2O of Water bears TIVON to Water).

¹⁸⁷ For current purposes I take the difference between Chemical and Material to be that there are physical objects that are not, nor are they made of, chemicals (such as those subatomic particles that make up the elements). However, the specifics of the example are not of importance, and the reader is free to substitute in more palatable example as they wish.

Take ‘Primary Explanatory Feature’ to be the category of whatever feature(s) play the role of P-property, and are thus considered to be definitive of a kind; that feature by which it is distinguished from other kinds of the same level, and from which we get that kind’s membership criteria. The more general the kind is the more difficult it is to specify a proper name for Primary Explanatory Feature. Primary Explanatory Feature is the kind of property of which Elemental Composition, and Biological Origin are instances; Primary Explanatory Feature is simply the most general category of this kind. That these properties are primary explanatory features for kinds is a robust fact about those kinds, but the significance of primary explanatory features to the definition and membership criteria of a kind are a matter of convention.

The predication of Elemental Composition may at first look strange. To say that Chemical essentially has Elemental Composition, whilst true, seems trivial and in no way separates chemical kinds from any other kinds (presumably all physical objects essentially have Elemental Composition, even if they do not essentially have a specific elemental composition). This confusion is the fault of the phrasing of the GPI (which was maintained for the sake of simplicity and maintaining similarity with Sidelle’s GPIs). The role that Elemental Composition is playing here is that of a value of Primary Explanatory Feature, just as H_2O is the value of Elemental Composition. So when MATERIAL asserts that Elemental Composition is essential to Chemical it means that Elemental Composition is essentially the Primary Explanatory Feature of Chemical. With this clarified, MATERIAL might be rephrased as:

MATERIAL: If Chemical belongs to Material, then (if Elemental Composition is Chemical’s Primary Explanatory Feature, then the proposition that predicates having Elemental Composition as the Primary Explanatory Feature of Chemical bears TIVON to Chemical).

We might then introduce the next link in the chain: a GPI for individual quantities of water:

WATER: (x) (If x belongs to Water, then (if H_2O is x 's Elemental Composition, then the proposition that predicates H_2O of x bears TIVON to x)).

WATER furnishes us not only with the membership criteria for the kind water, but provides us with the grounds for individual essentialist facts about objects. A particular quantity of water has essential properties in accordance with the P-properties of the kinds it is a member of. The same can be done for objects like Socrates, for whom the relevant GPI might be:

SOCRATES: If Socrates belongs to Animal, then (if Phaenarete-and-Sophoniscus is Socrates' Biological Origin, then the proposition that predicates Phaenarete-and-Sophoniscus of Socrates bears TIVON to Socrates).

Where 'Phaenarete-and-Sophoniscus' is short for a description of Socrates' biological origins, and where SOCRATES is nested within a suitable hierarchy going up into more and more general kinds.

It is worth noting that there is a difference when we move from applying the mechanism to give us the essences of kinds, to using it for the essences of specific instances of those kinds. Using the Sidelleian mechanism there are two kinds of properties that objects can have essentially in virtue of the kinds they belong to. These are made clear by the examples used in Sidelle's work, and in this project. On the one hand, water essentially has the elemental composition H_2O because it is a chemical kind and the GPI governing water shows us that if water is a chemical kind, and if H_2O is its elemental composition (the P-property of chemical kinds) then water is essentially H_2O . On the other hand, Socrates is essentially the son of Phaenarete and Sophroniscus because he is a human and, as a human, whatever biological origins Socrates has, he has them essentially.

Prima facie these both appear to be instances of something being of a kind, and whatever property that thing has that is the value of the kind's P-property (e.g. H_2O is a value of elemental composition), it has essentially. This secures that Socrates essentially has the biological origins that he does (whatever that may be) and water

has the elemental composition that it does (whatever that may be). However, what is not explicitly accounted for is how specific quantities of water are essentially H₂O. This seems to fall through a gap that results from confusion between how GPIs govern kinds and individuals differently.

I suggest we distinguish between P-properties and *P-categories*. P-properties are properties of objects and kinds, whereas P-categories are categories of properties, and associated with kinds alone. P-properties are properties that something has that fall under a P-category. A P-category is a type of property, of which the P-properties that fall under it are potential values. For example, Socrates is the son of Phaenarete and Sophroniscus; this is his biological origins; call this specific biological origin O. O is a P-property of Socrates because Biological Origins (the category of property of which O is an instance) is a P-category of the kind Human, of which Socrates is a member.

Some kinds have both P-properties and P-categories. If kind K (to which x belongs) has P-property P, then so does x . If K has P-category Q then whatever property R x has that is a value of Q is a P-property of x . Object x is essentially F iff x is F, and F is a P-property of x . F is a P-property of x iff F is a P-property of a kind to which x belongs, or F is a value of a P-category of a kind to which x belongs. For example, Eddie the dog is essentially a mammal, because being a mammal is a P-property of the kind Dog; Eddie also essentially has whatever his biological origins are because they are the value of a P-category of Dog (biological origins).

Belonging to a kind is a relation between a thing and a kind it falls under (e.g. Socrates belongs to the kind Human), or between a subkind and a superkind (e.g. the kind Human and the kind Animal). Belonging to is a transitive relation, so if Socrates belongs to Human, and Human belongs to Animal, then Socrates belongs to Animal. As such, P-properties of Mammal (e.g. being warm blooded) or Human (e.g. being bipedal) are P-properties of Socrates, and properties that Socrates has that are values of the P-categories of Animal (e.g. species) or Human (e.g. biological origins) are P-properties of Socrates.¹⁸⁸

¹⁸⁸ Note that belonging to is not a relation between kinds that fall under each other as described earlier in Section Five. This ensures that just because Socrates belongs to the kind Human, and Human falls under the kind Kind, Socrates does not inherit any of the P-properties of the kind Kind.

5.2 Robust-Objects or No-Objects?

With very little work the proposed account can be made to fit with either the no-TIVON no-objects option, or the no-TIVON robust-objects option. If the account is matched with the no-TIVON robust-objects option then a consequence of this position is a change in focus from the pre-theoretical understanding by which neo-Aristotelian essentialism was initially characterised. The primary locution given by Fine as the quintessential essentialist qualifying phrase is ‘... if it is to be the thing that it is’. Under the current position, the essence of an object has nothing to do with it being the thing it is, but rather with it being the *kind of thing* that it is. If objects are robust, but essentialist facts are not, then the essence of an object is a consequence of the kinds that it belongs to, and some facts about *how* that object is.

The account is compatible with both the no-TIVON robust-objects option, and the no-TIVON no-objects option. If it is paired with the no-TIVON no-objects option then essence remains concerned with a thing being the thing that it is, as well as being the kind of thing it is. This is more in line with the motivations behind Fine’s neo-Aristotelian essentialism. As such, I shall proceed focusing primarily on the no-TIVON no-objects option (though what is said is compatible with the no-TIVON robust-objects option as well).

5.3 Dispensing with the Propositional Formulation of Essentialist Predication

With an antirealist theory of neo-Aristotelian essence established, the next task is to dispense with the propositional characterisation of essence and the mysterious TIVON relation. In the realist theory of essence provided by Fine, TIVON was taken to be an unanalysed relation between propositions and objects.¹⁸⁹ In the antirealist theory under consideration in this chapter, the conventional nature of the Sidelleian mechanism requires that we give a conventional explication of TIVON, explaining

¹⁸⁹ Fine (1995b), p.273

exactly when it holds between a proposition and an object. With an explanation given, including TIVON in the theory becomes somewhat redundant. The relation is a vestigial remnant of its realist counterpart and the current theory is better served by a more explicit formulation that does not rely on it.

If for a to be essentially F is for the proposition $\langle Fa \rangle$ to bear TIVON to a , and for $\langle Fa \rangle$ to bear TIVON to a is just for it to be the case that we do not consider anything to be a that is not F , then for a to be essentially F just is for it to be the case that we do not consider anything to be a that is not F . This is well served by the Sidellean GPI conventional schema:

GPI Conventional Schema: If ' x ' is a K -term then if p is the P -property of the thing denoted by ' x ', then ' x ' applies to something in any possible situation only if it is (has) p .

And the object language GPI schema can take the form:

GPI Essentialist Schema: (x) (If x belongs to kind K , then (if p is x 's P -property, then x is essentially p))

Dispensing with the propositional formulation and TIVON in this way reveals a potential concern that one might have with antirealist neo-Aristotelian essentialism as presented in this chapter. The suspicion is that the account is in fact not one of neo-Aristotelian essence, but of modal essence in a thin disguise. The similarity between the GPIs presented in this chapter and those presented by Sidelle is striking, especially with TIVON removed. By replacing 'necessarily' with 'essentially' in the GPI, do we really escape from modal essentialism, especially when the modal essentialist takes a to be essentially F iff a is necessarily F ?

The answer to this question is both 'yes' and 'no'. Just as in Section Four I stressed that Sidellean conventionalism about modal essence bridges the gap between modal and neo-Aristotelian essentialism, so too does Sidellean conventionalism about neo-Aristotelian essence. The key difference between the two positions is in what they give priority to. The conventionalist about modal essence takes the essential facts about an object or kind to be those that are necessary of that object or

kind; those facts that are the conclusion of an inference from certain facts and GPIs. The conventionalist about neo-Aristotelian essence takes the essentialist facts to be those that follow from things being what they are (be that the object they are, or the kind of object they are).

On the modal essentialist account the objects and kinds are carved using the conventions that give us necessity. These necessities are the essentialist facts. Our conventions about under what conditions we apply object and kind terms ground the modal facts, but are themselves not modal. Under the neo-Aristotelian account those same conventions are not taken to directly ground the modal facts, but rather the essentialist facts. By giving us the objects and kinds, those conventions provide us with propositions that bear TIVON to those things. If the Finean essentialist theory of modality is to be accepted (see Chapter Five), then necessity does not come directly from the GPIs, but from the essentialist facts they ground.

The neo-Aristotelian essentialist takes the essential properties of an object to be those that it has in virtue of it being the thing that it is. This is the intuitive content that was supposed to back up our understanding of the unanalysed TIVON relation. However, there is no sense in which *a* is F because *a* is what it is. The only reason that *a* is F is because of the way the world is.^{190, 191} *What a is*, as understood here, is only responsible for giving certain of *a*'s properties a certain significant status, the status of being essential to *a*.

However, remember the primary locution from Fine. This claimed that *a* must be F if it is to be the thing that it is. The focus on the qualifying phrase at the end distracts us from the importance of the 'must' at the beginning. This phrase should not be taken as indicating that the object under consideration is F because it is the thing that it is, but that *it must be* F because it is the thing that it is. This need not be interpreted as claiming that *a*'s being what it is makes it the case that it is F, as the

¹⁹⁰ One might claim that if *a* is not a robust object, and rather something carved out of the world by us, then *a* is F because of what it is, because if we had carved differently *a* might not have been F. However, this would be true of all of *a*'s properties, and so this response goes too far, making all of *a*'s properties essential to it.

¹⁹¹ Prima facie, *a* being F because of the way the world is, and *a* being F because *a* is what it is, do not sound mutually exclusive. After all, *a* is a part of the world. However, the way the world is, under the current position, does not specify *what a is*, only *how a is*.

weaker interpretation is available that it is the importance of *a* being *F* that is due to *a*'s being what it is, not *a*'s being *F* in the first place. So whilst the current account does use the same mechanism as the modal essentialist account, it also fulfils the criteria that Fine sets out. It matches the pre-philosophical content of the primary essentialist locution; the essential properties of an object are determined by the object being what it is (whether that is it being the kind of object it is, or just being the object it is), and, on the assumption that the Sidellean account does not assume anything fundamentally modal, does not preclude the possibility of an essentialist theory of modality (see Chapter Five for further discussion of this point).

6. Generic Essence

So far there has been much casual talk of kinds and their relationships to each other. This might give the impression that the positions accept a sizable ontological commitment through reifying kinds. That is something I am eager to avoid.¹⁹² Here I aim to interpret talk of kinds as minimally as possible. I take the essentialist claims about kinds to be those expressible by Correia's generic essentialist claims.¹⁹³ I talk in terms of kinds and their essences for convenience and the reader's ease, but this talk should not be taken as committing one to any entities that are kinds, but rather should be taken as translatable into the language of generic essence.

In this section I present Correia's distinction between objectual and generic essence, along with his arguments against plausible accounts of generic essence in terms of objectual essence. In doing so I show that adoption of generic essentialist claims renders two advantages to the antirealist about essence over the realist. First, interpreting the essences of kinds in terms of generic essence deflates the *prima facie* ontological obligations of the theory. Second, I argue that the antirealist is better placed to incorporate generic essence into their account than their realist counterparts. Just how the antirealist accounts for generic essence depends on

¹⁹² I am by no means alone in this. See Shalkowski (forthcoming).

¹⁹³ Correia (2006)

whether one adopts the no-TIVON no-objects option or the no-TIVON robust-object option.

Correia distinguishes between objectual and generic essentialist statements. *Objectual essentialist claims* tell us what it is to be a specific object; they take the familiar form ‘*a* essentially Fs’. *Generic essentialist claims* tell us what it is to be of a certain kind, or to have a certain property; they take the form ‘Fs, as such, essentially G’. As Correia points out, the distinction between objectual and generic essence runs parallel to the distinction between subjects and predicates in language. Objectual essentialist claims describe what is essential to the referents of subject terms, namely *things*. Generic essentialist claims describe what is essential to that which predicates express, namely *ways of being*.¹⁹⁴

Correia claims that in the contemporary literature the prevailing attitude is either to ignore generic essentialist claims, or to view them as a special case of objectual essence, perhaps as the objectual essences of general entities such as properties or kinds, or as the objectual essences of all of the individual objects that have that property/are of that kind. In response to this Correia argues that generic essentialist claims cannot be characterised in terms of objectual essence, and that essentialism is best understood in terms of both objectual and generic essence. He does this by considering the potential accounts of generic essence in terms of objectual essence (on both the modal and neo-Aristotelian understanding of essence) and objecting to each in turn.

If Correia’s arguments are successful, then this means that talk of the essence of kinds should be understood in terms of generic essentialist claims. This in turn means that any theory of essence that wants to maintain essentialist claims about kinds will have to be able to accommodate generic essence. I take Correia’s arguments to be successful, and argue that realist neo-Aristotelian essentialism cannot accommodate generic essentialist claims as well as the antirealist position presented here. I take this advantage to be a significant motivation for antirealist essentialism.

¹⁹⁴ Correia (2006), p.754

Here I present a brief summary of Correia's arguments against the various options for explaining generic essence in other terms.¹⁹⁵

Option 1

Under the modal conception of essence one might extend the account of objectual essence to explain generic essence by claiming that the generic essentialist claim 'Fs, as such, are essentially G' is true iff it is essential to all of the Fs that they are G. That is:

MODAL: Fs, as such, are essentially G iff necessarily, every F is G.

However, such an account is vulnerable to the same Finean criticisms of the modal account of objectual essence. Every person is such that they belong to the singleton set of just themselves. It follows from MODAL that people, as such, are essentially members of their own singleton sets. One may wish to adapt the account so as to exclude universal features, like belonging to one's own singleton, like so:

MODAL 2: Fs, as such, are essentially G iff necessarily, every F is G, and possibly, something is not G.

However, there are still problematic set memberships. Not everything is a member of the set of people, yet all of the people are. MODAL 2 does nothing to preclude that people, as such, are essentially members of the set of all people.

Option 2

The second option explains generic essence in terms of objectual essence and necessity.

OPTION 2: Fs, as such, are essentially G iff necessarily, every F is essentially G.¹⁹⁶

¹⁹⁵ Correia (2006)

¹⁹⁶ Presumably the neo-Aristotelian conception of objectual essence is intended here. If the modal conception were intended then the analysis would be equivalent to claiming: Fs, as such, are essentially G iff necessarily, every F is necessarily G. Such an account would not escape the problems of the modal account, and Correia takes this account to do just that.

However this account does not accommodate all of the essentialist claims that it should. Bachelors, as such, are essentially unmarried, but for that to be the case under the current account would require all of the individual bachelors to be essentially unmarried. It would be the unlucky bachelor indeed who was essentially unmarried, so this account should also be rejected.

Option 3

Option 3 involves the use of *qua-objects*.

OPTION 3: Fs, as such, are essentially G iff necessarily, for every x such that x is F, x -qua-F is essentially G.

This account successfully avoids the problems faced by the previous options, because the problematic features of the objects in question are not features that they have qua-F. That Socrates is a member of the singleton set of himself is not essential to him, and even the luckiest bachelor is essentially unmarried qua-bachelor.

However, this account must also be rejected. Ignoring whatever consequences may come from adopting the use of qua-objects, Correia identifies three problems. First, because there can be no round squares, according to the qua-object account round squares can have any property essentially (because the antecedent of the conditional in the second half of OPTION 3 is necessarily false). However, it is reasonable to think that if round squares are essentially anything, they are essentially round and not (say) essentially bready. Second, any qua-object will be essentially a qua-object. However, Correia claims, to think that humans are essentially qua-objects is absurd. Third, the account attributes explanation of essential predication in the wrong direction. That bachelors, as such, are essentially unmarried is explained by necessarily, all of the individual bachelors being qua-bachelor essentially unmarried. However, intuitively it should be because bachelors, as such, are essentially married, that the individual qua-bachelor objects are essentially unmarried. As such, Correia rejects this option as well.

Option 4

Instead of generic essences being explained in terms of the objectual essences of the members of kinds, perhaps it should be explained in terms of the objectual essences of those kinds or properties themselves. This account claims that:

PROPERTIES: Fs, as such, are essentially G iff it is true in virtue of what the property of being F is that Fs are G.

Whilst Correia shows that this account does not face any of the problems faced by the above mentioned alternatives, he gives three problems what he takes to be sufficient to motivate the rejection of PROPERTIES.

First, the general form of generic essentialist claims does not appear to commit us to the reification of properties. It does not seem to be contradictory to hold the view that there are true generic essentialist claims, but no properties. However, under this proposal there is such a commitment. This is not a devastating objection, but for those reluctant to endorse an ontology of properties it may be a serious deterrent.¹⁹⁷ Second, some predicates cannot express properties, and yet can be the subject of true generic essentialist claims. For example, the predicate ‘... is a non-self-exemplifying property’ cannot have a corresponding property because for there to be such a property would lead to paradox regarding whether it self-exemplifies. However, there are true generic essentialist claims to be made that feature such a predicate. For example, ‘non-self-exemplifying properties, as such, are properties’ is a true generic essentialist claim. Third, the account, along with the plausible claims:

(M) It is true in virtue of what the property of being a man is that every man exemplifies it.

(T) If *a* is essentially an F, and if an F, as such, is essentially G, then *a* is essentially G.¹⁹⁸

¹⁹⁷ Rejecting properties in this sense is not to claim that ordinary predications, such as ‘Socrates is bearded’ are false, but to deny that the truth of such predications relies on the existence of reified properties.

¹⁹⁸ Correia (2006), p.762

commits us to the ontological dependence of Socrates (a man) upon the property of being a man. This is because if Socrates is essentially a man, then he essentially exemplifies the property of being a man. If there were no such property, then there could be no Socrates. However, for the same reasons that we should be reluctant to think that Socrates is essentially a member of certain sets, we should be reluctant to think that our theory of essence should commit us to Socrates being ontologically dependent on certain properties. Correia claims that even assuming the truth of (M) and (T) should not commit us to the claim that Socrates is ontologically dependent on the property of being a man, or of any property.

6.1 Realism and Generic Essence

If Correia's arguments are successful then this leaves the essentialist in the position where talk of what up until now we have been calling the essences of kinds is best understood in terms of generic essentialist claims. This means that one must accept generic essence as being at least as basic as objectual essence. This has different consequences for the modal and neo-Aristotelian realist essentialist.

Correia's arguments against the modal account of generic essence work by showing the extensional inaccuracy that results from understanding generic essence in terms of the objectual essences of all of the members of the kind in question. In this way Correia's arguments against the modal account of generic essence work in the same way as Fine's arguments against modal objectual essence. However, the modal essentialist appears to be in no worse a position than they were previously. Correia's arguments add little to the case against modal essentialism beyond that which has already been made. They take the same form as the criticisms Fine gives of modal essentialism in general. As such, how decisive one takes Correia's arguments to be depends largely on how one views Fine's arguments. If Fine's criticisms can be overcome, and thus modal essentialism is a viable position to start with, then it is not clear that Correia's criticisms still hold. If Correia's arguments do still hold, then so

do Fine's criticisms, and the modal essentialist arguably has more pressing concerns than accounting for generic essence.

However, Correia's arguments do present a significant challenge for the Finean neo-Aristotelian realist essentialist. If essence is explained in terms of TIVON, and TIVON is a relation between propositions and objects (broadly construed so that it could include properties, etc.), then presumably the Finean neo-Aristotelian account of the essences of kinds would be to say that Fs are essentially G iff the proposition <Fs are G> bears TIVON to the property/kind/concept of being F. For example, that bachelors are essentially unmarried can allegedly be explained as:

BACHELORS: Bachelors, as such, are essentially unmarried iff <Bachelors are unmarried> bears TIVON to Bachelorhood (where 'Bachelorhood' is either the property of being a bachelor, the kind Bachelor, or the concept Bachelor).

If the claims we want to make about the essences of kinds are identical to those expressed by the generic essentialist claims, then BACHELORS encounters the same problems as PROPERTIES. The essences of kinds cannot be characterised in terms of the objectual essences of kinds/properties, but rather must be understood as generic essentialist claims. As such, the Finean explanation is inadequate, and generic essence must be adopted as primitive.

However, this too is problematic for the realist. Generic essentialist claims cannot be explained in terms of TIVON because there are no appropriate relata for propositions to bear TIVON to (if TIVON obtained between the proposition and the property or kind then it encounters the same problem as PROPERTIES, if it bears TIVON to the objects that belong to the kind then it encounters the same problem as the previous failed options). There is no obvious alternative explanation that might be given in its place, save that generic essence is primitive. The resulting theory, however, is one where objectual essence has further explanation (in the primitive TIVON relation), but generic essence does not, which is odd.

One might opt to abandon the TIVON based explanation of objectual essence, and merely claim that essence (both objectual and generic) is primitive, perhaps relying on talk of TIVON only as a useful metaphor that serves as a heuristic for the

account. However, this would still leave the essentialist in a position where there are two primitive notions of essence. At this point the account starts to look rather inelegant. A theory with multiple primitives of so similar a nature is worthy of suspicion.

There is a further alternative that is available to the realist. Correia suggests that we explain objectual essence in terms of generic essence.¹⁹⁹ For each object a there is a corresponding haecceity ϕ , a property of what it is to be that thing. As such we can say that any objectual essentialist claim ' a is essentially F ', which is understood as 'it is true in virtue of what a is that Fa ' can be reduced to a corresponding generic essentialist claim to the effect that it is true in virtue of the haecceity ϕ that Fa . This certainly escapes the discomfort of multiple similar primitives, but the reliance on haecceities is somewhat troubling.

Ignoring whatever else might be controversial about haecceities, to make use of them in this way seems to skirt the issue. If the claim is that the haecceity is the property of what it is to be a certain thing, then it assumes that there is something that it is to be that thing. If this is the case then how can this not be the objectual essence of the thing at hand? Moving from the objectual to the generic was an acceptable move in the case of regular predicates because there was no other candidate for an object for the essence to be of, but to do so when there is the perfect candidate right there almost seems to be sleight of hand. Whilst the reduction works, it does so in spite of a more obvious position.

6.2 Antirealism and Generic Essence

The antirealist account has the advantage here, as using the conventionalist mechanism developed from Sidelle, the antirealist about neo-Aristotelian essence can account for generic essence using the same tools they use to account for objectual essence. Thus, the antirealist account is unified in a way that the realist account is

¹⁹⁹ Correia (2006), p.765

not. Furthermore, adopting generic essence as the proper formulation of essentialist claims about kinds enables the antirealist to avoid ontological commitment to kinds.

To avoid this commitment we must be able to show what it is for an object to be of a kind. For a to be of kind F is just for F to be truly predicated of a . It is also important to show that subkind/superkind claims can be made using just generic essence. This can be done: one kind F is a subkind of a more general kind G iff for any x , if x is F then x is G (that is, if F can be truly predicated of x then G can be truly predicated of x). All that remains to be shown is that the Sidellean mechanism can be made to work for generic essentialist claims, understood as such, and not as objectual essentialist claims about kinds.

Just as with the Sidellean account of objectual essence, the generic essences of kinds (that is, the essences of kinds, expressed as generic essentialist claims) are determined by the superkinds to which they belong. Socrates is essentially the son of Phaenarete and Sophroniscus because he is a human; because being the son of Phaenarete and Sophroniscus is his biological origins; and because biological origin is the P -category for humans. Likewise, generic essentialist claims are those claims that are the result of a valid inference from an appropriate GPI and an additional claim. That F s, as such, are essentially G is true because F s are G , and because being G is the P -property (or the value of a P -category) for kind- K , which is a superkind of F . We get the account:

GENERIC Account: <An F , as such, is essentially G > is true iff it is the conclusion of a valid inference from an additional claim and an appropriate GPI.

This GPI is established by the following conventional practice:

CONVENTIONAL GENERIC GPI: If ' F ' is a K -term then if G is the P -property of the things denoted by ' F ', then ' F ' applies to something in any possible situation only if it is G .

This can be applied to specific cases. The generic claim that water, as such, is essentially H_2O , is true because all/enough of the things that we apply ‘water’ to are H_2O , and because we hold the following GPI:

GENERIC WATER: (x) (if water is a chemical kind, then (if H_2O is water’s elemental composition, then water is essentially H_2O))

This GPI is established by the following conventional practice:

GENERIC WATER’: If ‘water’ is a chemical kind-term then if H_2O is the elemental composition (the P-category) of the things denoted by ‘water’, then ‘water’ applies to something in any possible situation only if it has the elemental composition H_2O .

This is sufficient to establish the truth of the generic essentialist claims. Furthermore, this is done with recourse only to conventions that are already used in the theory. The GPIs that give us the generic essentialist claims are the same conventions used before the introduction of the objectual/generic distinction in the first place.

7. Conclusion

In this chapter I explored the various options available to the antirealist about neo-Aristotelian essence. Using the modified Sidelleian conventionalist mechanism established in Chapter Two I developed two of these options into viable accounts and showed how they can account for neo-Aristotelian essentialist claims without having to resort to the mysterious TIVON relation. I then used Correia’s work to argue that the essences of kinds are better understood in terms of generic essence than in terms of the essences of individual objects, and used this conclusion to motivate the positions developed in this chapter over their realist counterparts. The inability of the realist neo-Aristotelian essentialist to present a unified theory of objectual and generic essence leaves the position compromised in a way that the antirealist position is not.

In the next chapter I examine different attempts at a neo-Aristotelian essentialist theory of necessity. I evaluate and provide commentary on these theories, and then explore the prospects of an antirealist appropriation of the realist strategy to see if the antirealist essentialist can do the same.

Chapter Five: Essentialist Theories of Necessity

1. Introduction

The primary motivation for adopting neo-Aristotelian essentialism over modal essentialism comes from neo-Aristotelian essence's superior ability to integrate into our wider metaphysics. Whilst the arguments given against modal essentialism were shown not to be decisive, failing to demonstrate that it doesn't make sense to understand essentialist claims in terms of modality, it is evident that understanding essence modally does nothing to help us understand the world any more than we already did. After all, under modal essentialism, to know the modal facts is to know the essentialist facts; identifying the subset of modal facts that are the essentialist facts tells us nothing about the world except what we mean when we say "... is essentially ...". However, interpreting essentialist claims as the neo-Aristotelian does allows us to give essence an explanatory role in our wider metaphysics. An important part of this is the assertion that using neo-Aristotelian essentialist facts we can give an explanation of metaphysical necessity. The claim that this can be done is well established.²⁰⁰ However, only a few have made a serious attempt at providing a developed account of this explanation. In this chapter I present and evaluate the options for a realist neo-Aristotelian essentialist theory of necessity. Having established the virtues and challenges of such accounts, I determine whether such a strategy can be appropriated by the antirealist as presented in Chapter Four. I conclude that not only can the antirealist give such a theory of necessity, but that the antirealist essentialist theory of necessity has certain advantages over its realist counterparts.

In Sections Two, Three, and Four I present, compare, and develop the realist essentialist theories of necessity as developed (separately) by Fine, Correia, and

²⁰⁰ Cameron (2010), Correia (2006, 2012), Fine (1994, 1995c), Hale (2013), Lowe (2008), Shalkowski (2008).

Hale.²⁰¹ There are key differences in their theories of essence and of modality, and I examine these. In doing so I present potential problems with those accounts and explore how they might be overcome. In Section Five I conclude by presenting the advantages and disadvantages of each position, showing that each has their flaws. It is against these conclusions that any antirealist essentialist theory of necessity is to be judged. In Section Six I develop an antirealist essentialist theory of necessity. I argue that it is a viable position, and has certain advantages over each of its realist competitors.

2. Fine

2.1 Finean Essentialism

In Chapter Three, we settled on a formulation of neo-Aristotelian essence of the form:

ESSENCE: *a* is essentially *F* iff the proposition $\langle Fa \rangle$ bears TIVON to *a*, and *a* alone.

where TIVON is a unanalysed relation between propositions and objects.²⁰² The essence of an object is the collection of propositions that bear TIVON to that object.

TIVON is the key primitive of the theory; essentialist facts are grounded entirely in instances of TIVON holding between propositions and objects. A proposition bears TIVON to an object (or group of objects) when that proposition is true because those objects are the objects that they are. The propositions that bear TIVON to an object form a subset of the propositions that are made true by that object. Whilst $\langle \text{Socrates had a beard} \rangle$ may be true because Socrates existed and had a beard, that proposition does not bear TIVON to Socrates. However, $\langle \text{Socrates was a human} \rangle$ does bear TIVON to Socrates because there is something about Socrates being

²⁰¹ Fine (1994, 1995a, 1995b, 1995c, 2000), Hale (2013), Correia (2006, 2012).

²⁰² The TIVON relation is a part of the machinery of Fine's account, but as we will see in this chapter there is more than one way to interpret the 'true in virtue of the nature of' locution. As such, when I use 'true in virtue of the nature of' instead of TIVON, I intend this to be interpreted neutrally between different interpretations, or trust context to make it evident which is intended.

Socrates (and not just existing) in particular that makes <Socrates was a human> true.

2.2 The Finean Essentialist Theory of Modality

Fine argues that rather than essence being a special case of metaphysical necessity, the converse is true and metaphysical necessity is a special case of essence.²⁰³ Chapter Three saw the arguments Fine gave for his negative claim; here I present the positive claim.

Fine identifies that “any essentialist attribution will give rise to a necessary truth”.²⁰⁴ If this is the case then it is only a small step to claim that:

[E]ach class of objects, be they concepts or individuals or entities of some other kind, will give rise to its own domain of necessary truths, the truths which flow from the nature of the objects in question.²⁰⁵

As such, the essences of the logical concepts will give rise to the class of logical necessities, the essences of numbers and mathematical concepts will give rise to the class of mathematical necessities, etc.²⁰⁶ Ultimately, “[t]he metaphysically necessary truths can then be identified with the propositions which are true in virtue of the nature of all objects whatever.”²⁰⁷ This has the consequence that all forms of necessity are special cases of metaphysical necessity. If one can provide an explanation for all of the true necessitation claims then, through the interdefinability of necessity and possibility, one can provide a full account of the basic modal notions.²⁰⁸ As such, if metaphysical necessity is a special case of essence, and other forms of necessity are special cases of metaphysical necessity (because those forms of necessity are the

²⁰³ Fine (1994), p.8

²⁰⁴ Fine (1994), p.7

²⁰⁵ Fine (1994), p.8

²⁰⁶ Talk of concepts here comes straight from Fine. I reserve scepticism about talk of concepts for later in this chapter. For now I present Fine’s position as he does.

²⁰⁷ Fine (1994), p.8

²⁰⁸ One needs only one modal notion (and negation) in order to define the others. For example, necessarily P iff not-possibly-not P; possibly P iff not-necessarily-not P; and it is impossible that P iff necessarily-not P (or not-possibly P).

necessities that objects of the relevant domain give rise to), then by showing that metaphysical necessity can be explained in terms of essence, Fine gives an account of modality.

These preliminary remarks show how the essentialist theory of necessity is to proceed. The claim is that for every necessitation claim $\text{nec}(P)$ there is an essentialist claim that grounds that necessity. A proposition P is necessarily true iff P bears TIVON to some thing or things. However, what these remarks do not establish is that all necessity has its grounds in the essences of things. This is the next task here.

2.3 Finean Logic of Essence

To demonstrate the expressive and extensional adequacy of his essentialist theory of necessity, Fine develops a logic of essence. Fine introduces an essentialist operator that replaces any modal operators in the logic. Read $\Box_F A$ as the claim that sentence A is true in virtue of the nature of (some or all of) the F s (the things that satisfy the predicate F). F in this instance is called the delimiter. If $\Box_F A$ then we can say that the proposition expressed by A bears TIVON to (some or all of) the F s. The language uses possibilist quantifiers. Take Π and Σ as universal and existential quantifiers (respectively) with the domain of all possibilia. For ease of presentation, sometimes the essentialist operator appears relativized to a name or variable. Note that in these situations the operator is in fact relativized to a predicate that singles out that name or variable. These are called rigid predicates; predicates of the form ‘... is either a , or b , or c ...’ for $n \geq 1$ objects. For instance $\Box_a Fa =_{\text{def}} \Box_G Fa$ (where G is the rigid predicate that singles out a), and $\Sigma x \Box_x Fx =_{\text{def}} \Sigma x \Box_{(\lambda y)(y=x)} Fx$.

According to the preliminary essentialist theory of modality, a proposition is necessary when it bears TIVON to (some or all of) the members of the class of all things. Read $\Box_V A$ as the claim that A is true in virtue of the natures of (some or all of) the members of the class of all objects. The delimiter V is a predicate that is universally satisfied. $\Box_V A =_{\text{def}} \Box_{(\lambda x)} A$. That is, A is true in virtue of the nature of (one

or more of the members of) the class of things such that T is true, where T is a sentence that is always true (as such, it is a predicate that everything satisfies).

Fine also introduces a secondary form of necessity. Read $\Box_{\wedge}A$ as the claim that A is true regardless of the nature of any objects (call this minimal necessity, as opposed to the maximal necessity expressed by $\Box_{\vee}A$). $\Box_{\wedge}A =_{\text{def}} \Box_{(\lambda x)\perp}A$. That is, A is true in virtue of the nature of (one or more of the members of) the class of things such that \perp is true, where \perp is a sentence that is always false (as such, it is a predicate that nothing satisfies). Fine takes the propositions expressed by a sentence A such that $\Box_{\wedge}A$ to be the logical or conceptual necessities.²⁰⁹ This point is of significance, and will be returned to in Section 2.9.

Fine develops an axiomatised logical system of essence equivalent to S5 in modal logic. He calls this system E5. He sets out to prove that for any theorem of S5 π (a quantified modal logic for S5 with a constant non-empty domain) there is a corresponding theorem of E5⁺ (the E5 system, strengthened with the domain axioms that (i) the existence of each object in the domain is compatible with the natures of every other object, and (ii) it is true in virtue of the natures of all of the objects that they are all of the objects) resulting from the substitution of each occurrence of ' \Box ' (the standard necessity operator) with ' \Box_{\vee} ' (Fine's maximal necessity operator).²¹⁰

2.4 Correia on Essence and Modality

Correia argues that "the plausibility of Fine's view crucially requires that certain apparent explanatory links between essentialist facts be admitted and accounted for".²¹¹ In investigating how these links are to be best accounted for by the essentialist, Correia modifies Fine's account in such a way as to improve the theory of logical and conceptual necessity.

Correia takes it to be illuminating to consider collective essences. The neo-Aristotelian essentialist is committed to there being irreducibly collective essentialist

²⁰⁹ Fine (1995b), p.250-251

²¹⁰ Fine (1995b), p.267

²¹¹ Correia (2012), p.1

attributions. That is, essentialist attributions that are true in virtue of the nature of a plurality of objects without there being any one object's nature that they are true in virtue of. For example, Socrates is essentially distinct from the Eiffel Tower. This is not true in virtue of the nature of Socrates alone, nor of the Eiffel Tower alone, nor even of the identity relation alone. To claim that it is would expose one to Fine's criticisms of modal essentialism.^{212, 213} Rather, it is true in virtue of the nature of the plurality of Socrates, the Eiffel tower, and identity.²¹⁴

Use X to denote a plurality of $n \geq 1$ objects, and α as a singular term for propositions. When only one object a belongs to X , $a = X$. Y is a part of X iff all objects that belong to Y belong to X . Correia discusses essentialist facts in terms of the truth of propositions in virtue of the natures of pluralities (in the liberal sense that includes pluralities of one).

Correia claims the neo-Aristotelian essentialist should distinguish between that which is *basically* essential, and that which is *derivatively* essential. That is, claims that are true in virtue of the natures of some group of objects, and for which there is no further explanation in essentialist terms (call these basically essential), and those for which there is further explanation available (call these derivatively essential). Whilst the distinctness of Socrates and the Eiffel Tower may be a good example of basic essentiality, some collective essentialist attributions are more plausibly derivative. For example, propositions reached by the conjunction of multiple propositions, each of which is true in virtue of the nature of a distinct part of a plurality X . The proposition $\langle\langle$ Socrates is organic \rangle and \langle the Eiffel Tower is synthetic $\rangle\rangle$ is irreducibly collectively essential to X (where X is a plurality containing Socrates, the Eiffel Tower, and the concept conjunction), but further explanation can be given in essentialist terms. The proposition is true in virtue of the natures of Socrates, the Eiffel Tower, and the concept conjunction taken together, *because* it is true in virtue of the individual nature of Socrates that \langle Socrates is organic \rangle ; true in

²¹² Fine (1994)

²¹³ This is supported in Fine (1995b), p.255

²¹⁴ Whether it is true only in virtue of the nature of the plurality that has just those things as parts, or of other pluralities that contain those things but other parts as well, will be the topic of the following discussion.

virtue of the individual nature of the Eiffel Tower that <the Eiffel Tower is synthetic>; and true in virtue of the nature of the concept conjunction that when two propositions are true, so is the conjunction of those two propositions. As such, it is derivatively essential to *X* that Socrates is organic and the Eiffel Tower is synthetic because <<Socrates is organic> and <the Eiffel Tower is synthetic>> admits further explanation in essentialist terms, whereas a proposition that is basically essential, such as <Socrates is distinct from the Eiffel Tower> admits no further explanation than its being true in virtue of the natures of Socrates and the Eiffel tower taken together. As such, it is basically essential. The *basic nature* of *X* (shown as $\beta(X)$) is the plurality of propositions that are basically essential to any *Y* that is part of *X*. We can define basic and derivative essence like so:

BASIC: For a plurality *X* to be basically essentially *F* is for <*X* is *F* if *X* exists> to be true in virtue of the nature of *X*, where no further explanation can be given in essentialist terms.

DERIVATIVE: For a plurality *X* to be derivatively essentially *F* is for <*X* is *F* if *X* exists> to be true in virtue of the nature of *X*, where further explanation can be given in terms of the natures of parts of *X*.²¹⁵

Correia sets out to provide an account of the derivatively essential; in doing so he refines the theory of modality in terms of essence. Without an account of derivative essence Correia claims that essentialism would be incomplete, and less convincing.²¹⁶

2.5 Consequentialist Account

Correia argues that one cannot account for derivative essence in terms of Finean consequential essence. It cannot be accounted for in terms of the consequential essence characterised in the way I develop in Chapter Three either.

Fine's notion of consequential essence results from essence being closed under logical consequence. As such, the consequential essence of an object includes

²¹⁵ Correia (2012), p.5

²¹⁶ Correia (2012), p.5

propositions that bear TIVON to it, but also propositions that are logical consequences of propositions that bear TIVON to it. One might think that the derivative essence of an object is made up of those propositions. According to such an account, α is derivatively essential to X iff it is a logical consequence of $\beta(X)$, but does not belong to $\beta(X)$. That is, α is derivatively essential to X iff it is part of the consequential, but not the constitutive essence of X .

However, the consequential account has poor results when combined with the Finean theory of necessity. Any logical claim necessitated is a logical consequence of any plurality of propositions Δ . As such it will be a logical consequence of the basic nature of any plurality X . This means that every logical necessity will be part of the derivative essence of every plurality of objects whatsoever. This conflicts with Fine's claim that the essence of an object does not include that which does not pertain to the nature of the given object.

There is a potential line of response that one might give on behalf of the consequentialist account. Whilst it is true that the consequentialist notion of derivative essence presented here includes in the derivative essence of any X , all logical claims necessitated, this does not accurately correspond to how consequential essence is presented by Fine. In *Senses of Essence* Fine claims that the consequential essence of an object excludes propositions that can be generalised away because they are true of everything, and in *The Logic of Essence* Fine claims that "we do not allow the logical consequences in question to involve objects which do not pertain to the nature of the given objects".²¹⁷ Is it not the case then that the consequentialist can exclude the unwanted propositions from the derivative essence of X ? No. As we saw in Chapter Three such generalisation responses cannot exclude all of the problematic propositions.

What of the definition of consequential essence I developed in Chapter Three, whereby the consequential essence of an object is the class of propositions that bear TIVON to the pluralities that the object is a part of? The natural attempt would be to

²¹⁷ Fine (1995a), p.56-61, (1995b), p.242 (I will not occupy myself with exactly what Fine means by "we do not allow".)

follow the previous consequentialist attempt and define derivative essence as those propositions that bear TIVON to the pluralities an object is a part of, but that do not bear TIVON to that object alone (thus excluding the constitutive essence of the object).

This definition will not suffice either. Like Fine's consequential essence, my notion of consequential essence is also extensionally inaccurate for the notion Correia has in mind. The proposition $\langle\langle Fa \rangle \wedge \langle Fb \rangle \rangle$ bears TIVON to the plurality of just a , b , and \wedge .²¹⁸ As such it is part of the constitutive essence of that plurality. If this is the case then by the definition given here that proposition cannot be derivatively essential to that plurality. This is one of Correia's own examples of derivative essence, so this definition is clearly not suitable. We might think to fix the account by stipulating that for a proposition to be derivatively essential to a plurality it cannot be part of the basic essence (rather than the constitutive essence) of that plurality. This is not enough either, as it will be derivatively essential to the plurality of just a that $\langle\langle Fa \rangle \wedge \langle Fb \rangle \rangle$, because it is a part of the larger plurality of a , b , and \wedge . Worse, it will be derivatively essential to the plurality of just c that $\langle\langle Fa \rangle \wedge \langle Fb \rangle \rangle$, because c is a part of the larger plurality of a , b , c , and \wedge , to which $\langle\langle Fa \rangle \wedge \langle Fb \rangle \rangle$ bears TIVON.

2.6 Rule-Based Account

Correia proposes a rule-based account of derivative essence that avoids these problems. He begins by building on a suggestion made by Fine about the character of the essences of logical concepts.²¹⁹

Logical concepts have two kinds of basically essential feature, those that are properly logical, and those that are not. The properly logical features that are basically essential to a logical concept are inferential in character. They are not given in terms of propositions, but in terms of rules of inference. The properly logical

²¹⁸ By ' \wedge ' I mean what Fine calls the logical concept conjunction itself, not the connective that represents it in the language.

²¹⁹ Correia (2012), p.8, Fine (1995a), p.57-58

features in a logical concept's basic essence are those concerning the inferences that use of the concept licenses one to make. For instance, the concept disjunction essentially has the properly logical feature of licensing the inference of $P \vee Q$ from P . These inferential features should be distinguished from propositional features such as $\langle \text{disjunction is a concept} \rangle$, which are no less basically essential to disjunction, but not properly logical in character.

For his account of derivative essence, Correia develops a theory of relative logical consequence. Where S is a set of logical concepts and Δ is a plurality of propositions, α is a logical consequence of Δ relative to S (shown as $\Delta \vdash_S \alpha$) iff there is a proof of α from Δ using only rules licensed by members of S . Logical consequence simpliciter can be defined in terms of relative logical consequence like so:

$\Delta \vdash \alpha$ iff there is a set S of logical concepts such that $\Delta \vdash_S \alpha$.²²⁰

Using $\log(X)$ to represent the set of logical concepts that are part of X we can say that:

DERIVATIVE: α is derivatively essential to X iff α does not belong to $\beta(X)$, and
 $\beta(X) \vdash_{\log(X)} \alpha$.

From this we can then give the following account of essence:

ESSENTIAL: α is true in virtue of the nature of X iff $\beta(X) \vdash_{\log(X)} \alpha$.^{221, 222}

This means that logical necessities will only be derivatively essential to pluralities that contain the logical concepts involved in those propositions. This is a marked advantage over the first consequentialist account, according to which no plurality presents a better derivative essential ground for the truth of a logical necessity than any other. Under the consequential account the proposition $\langle \text{if pigs fly then pigs fly or dogs bark} \rangle$ is as derivatively essential to a plurality X with $\log(X) = \{\wedge, \exists\}$ as it is a plurality Y with $\log(Y) = \{\vee, \rightarrow\}$. This account also has the advantage over the second consequentialist account because $\langle \langle Fa \rangle \wedge \langle Fb \rangle \rangle$ is a logical consequence of the

²²⁰ Correia (2012), p.9

²²¹ Correia (2012), p.6

²²² Note that this does not constitute an analysis of 'true in virtue of the nature of', as the basic essence of X presumes that the proposition is already essential to X (and thus does nothing to explain how 'truth in virtue of the nature of' works).

plurality of a , b , and \wedge (relative to that plurality), and is not excluded, and it is neither a logical consequence of a nor c separately (relative to a and c respectively).

2.7 Logical and Conceptual Necessity

There is further advantage to the rule-based account that Correia proposes. The tools given in the exposition of the rule-based account allow the essentialist to refine their analysis of logical and conceptual necessity. The Finean reduction of metaphysical necessity, in Correia's terms, can be expressed as:

METAPHYSICAL: α is metaphysically necessary iff for some plurality X , $\beta(X)$

$\vdash \alpha$.²²³

Fine claims that a similar reduction holds for logical and conceptual necessity:

LOGICAL: α is logically necessary iff for some plurality X of logical concepts,

$\beta(X) \vdash \alpha$.

CONCEPTUAL: α is conceptually necessary iff for some plurality X of concepts, $\beta(X) \vdash \alpha$.

However, there is an additional problem for logical and conceptual necessity when it comes to moving from the weaker claim that logical concepts and concepts are the essentialist grounds for logical and conceptual claims necessitated, to the stronger claim made by the reduction given above, that *what it is* for a proposition to be logically necessary is for it to have such grounds. This is shown by the distinction between the basic essential features of a logical concept that are properly logical and those that are not. The proposition <if pigs fly then pigs fly or dogs bark> is logically necessary because it is true in virtue of the nature of a plurality containing the logical concepts disjunction and conditional. However, the proposition <disjunction is a concept> is also true in virtue of the nature of the same plurality, yet it is not a logical claim necessitated. The same problem arises for the reduction of conceptual necessity. There are essential features of concepts that result in non-conceptual

²²³ Correia (2012), p.11

necessities; for example <bachelorhood is a concept> is true in virtue of the nature of the concept bachelorhood, but is not a conceptual claim necessitated.

Correia works to resolve this problem by introducing the notion of relative theoremhood, corresponding to relative logical consequence as given above. A proposition α is a theorem relative to a set of logical concepts S iff $\vdash_S \alpha$. That is, if there is a proof of α from no premises, using only the logical concepts contained in S .

This furnishes us with an account of local necessity. The logical claims necessitated local to S are those that are theorems relative to S . From this we can provide an account of logical necessity simpliciter. A proposition is a theorem simpliciter iff it is a theorem relative to the set of all logical concepts. The logical claims necessitated are those that are theorems simpliciter.

Given a plurality X of logical concepts, the propositions such that $\beta(X) \vdash \alpha$ are divided into two mutually exclusive classes: those that are properly logical claims necessitated, and those that are not. The properly logical claims necessitated are those that are theorems relative to some set of logical concepts S . These propositions are true solely in virtue of the inferential nature of X , the properly logical essential features of the concepts it contains. The claims necessitated that are not properly logical are those that are true in virtue of the propositional nature of X , such as <disjunction is a concept>. Whilst still being (metaphysically) necessary, such propositions are not theorems and as such are not logically necessary. LOGICAL can then be re-defined as:

LOGICAL': α is logically necessary iff there is some set S of logical concepts such that $\vdash_S \alpha$.

A similar approach can be taken to modify CONCEPTUAL. For a plurality of concepts X , if it contains only logical concepts then the conceptual claims necessitated that are true in virtue of the nature of X are the theorems relative to X . If there are non-logical concepts in X then the basic nature of plurality Y of such concepts in X is divided between the properly conceptual propositions (such as <bachelors are male>) and the non-conceptual (such as <bachelorhood is a

concept>). The conceptual claims necessitated that are true in virtue of X then, are the logical consequences of the properly conceptual part of $\beta(Y)$ (written $\beta(Y)_C$) relative to $\log(X)$. We can re-define CONCEPTUAL as:

CONCEPTUAL': α is conceptually necessary iff for some plurality X of concepts, and plurality Y of non-logical concepts in X , $\beta(Y)_C \vdash_{\log(X)} \alpha$.

This escapes the problem as presented above in the same way that LOGICAL' escapes the problem for logical necessity. This has the advantage over the original reduction in that it avoids classifying <disjunction is a concept> as a logical claim necessitated, or <bachelorhood is a concept> as a conceptual claim necessitated. Presumably, for either of the adaptations to be successful, the notions of properly logical and properly conceptual are taken to be primitive, but this seems acceptable.

2.8 The Use of Concepts

Fine's theory of essence and the resulting theory of necessity both rely on concepts (both logical and non-logical) having essences. However, it is not clear what Fine means by his talk of concepts. Correia inherits talk of concepts from Fine, but does not expand upon what this might mean. Three questions arise. What does Fine mean by 'concepts'? Do we inherit any problems from using them? Must we rely on them?

The first two questions are difficult to give proper answers to. In answer to the first question Fine says almost nothing. One might be tempted to say that Fine is merely talking loosely, and by logical concepts what he really means are some form of platonic logical entity. Hale presents an essentialist theory of necessity using such logical entities.²²⁴ In what follows I consider Hale's position as an alternative to the use of concepts (and in doing so I give a negative answer to the third question). However, there is reason to think that this is not what Fine means by his use of concepts. When Fine introduces minimal necessity ($\Box \wedge A$, whereby A is true

²²⁴ Hale (2013)

regardless of the nature of any objects) he defines it as $\Box \wedge A =_{\text{def}} \Box_{(\lambda x)} \perp A$ (A is true in virtue of the class of things such that \perp is true, where \perp is a sentence that is always false, and thus not satisfied by anything). Fine takes the propositions expressed by a sentence A such that $\Box \wedge A$ to be the logical or conceptual necessities.²²⁵ If there are no objects such that a proposition can bear TIVON to them, then the concepts Fine speaks of as providing the source of logical and conceptual necessity must be something other than abstract objects. This tells us little about what Fine thinks concepts are, but suggests at least that he does not take them to be objects of some kind to which propositions can bear TIVON (which seems problematic in itself).

Without more of an answer to the first question, the second question is difficult to answer. However, it does seem that whatever position results from the use of concepts (construed as something other than objects) would be less unified than one that talks of logical objects that can be related to the TIVON relation just as chairs and people can. To have different forms of necessity accounted for in different ways detracts from the initial appeal of the essentialist theory of necessity.

In what follows I present Hale's essentialist theory of necessity. Hale's account does not rely on the use of concepts, and has some other significant differences from the position presented above. Once these various differences have been considered properly, in Section Six I present the final options for an account of necessity in terms of essence, against which the antirealist position will be compared.

3. Hale's Essentialist Theory of Necessity

Hale also presents an essentialist theory of necessity.²²⁶ However, Hale's explanation of modality in terms of essence is based on a theory of essence that, whilst neo-Aristotelian, differs in some important ways. Here I present Hale's theory of essence and his explanation of modality in terms of it. I identify the key differences between Hale's position and Fine's, and then adopt certain aspects of Hale's position so as to avoid problems encountered by Fine's position.

²²⁵ Fine (1995b), p.250-251

²²⁶ Hale (2013)

3.1 Hale on Essence

For Hale, the ‘true in virtue of the nature of ...’ locution represents a relation, not between propositions and objects as Fine claims, but between propositions and properties. For each thing there is a property that is the property of what it is to be that thing. This property is the *nature* of the thing in question. A proposition is true in virtue of the nature of an object when that proposition bears a relation to the property that is the nature of that thing. Whilst Hale identifies natures with properties, he claims to interpret them in a ‘metaphysically-lightweight’ or deflationary sense. Hale uses ‘nature’, ‘essence’, and ‘identity’ interchangeably with one another.²²⁷ In characterising the essence of a thing, Hale says:

By the nature or identity of a thing, I mean what it is to be that thing – what makes it the thing it is, and distinguishes it from every other thing. We may think of the nature or identity of a thing as what is given by its definition – that is, the definition of the *thing*, and not that of some word for the thing or concept of the thing.²²⁸

There are some interesting points to be drawn from this. First, Hale insists that the essence of an object should distinguish it from every other thing.²²⁹ This is an interesting departure from (at least, explicit) orthodoxy.²³⁰ It is not clear that the essence of a thing is always sufficient to distinguish something from every other possible thing. Hale gives an example of Aristotle essentially having his biological origins. Having his parents is obviously not sufficient to distinguish Aristotle from all of the other humans; possible siblings of Aristotle make for easy counterexamples. Even coming from the exact egg and sperm that Aristotle did is not sufficient; were Aristotle to have a monozygotic twin, they would act as a counterexample to this

²²⁷ Hale (2013), p.151, fn13

²²⁸ Hale (2013), p.132-133

²²⁹ Hale (2013), p.151

²³⁰ At least it is not the orthodoxy in the most recent literature. There is an older tradition (e.g. Plantinga (1974, 1976)) whereby the essence of an object must be necessary and sufficient to pick out that same object in every world in which it exists.

claim. The very possibility of such a twin is all that is needed for this counterexample. What about Aristotle's origins, or any part of his essence, could be sufficient to distinguish him from all other things? There are no obvious candidates in Aristotle's essence. Is it important that Aristotle be distinguished from every *possible* thing by his essence? Even if one were to answer 'no', the problem can be reconstructed using any pair of actual monozygotic twins. In light of this I will not pursue the claim that the essence of an object should distinguish it from every other thing. I do not think anything is lost in doing so.

Second, Hale does not identify the essence of a thing with the definition of that thing, as Fine does.²³¹ Rather, Hale identifies the essence of a thing with that which is given by that thing's definition. This has a significant consequence for the nature of the TIVON relation. According to Fine, TIVON is an unanalysed relation between propositions and objects.²³² Hale recognises this, but instead takes the 'true in virtue of the nature of ...' locution to be semantically complex, comprising a true-in-virtue-of relation, and the property that is the nature of the thing in question. This means that the mechanics of Hale's account do not rely on the mysterious TIVON relation, but on the more tractable true-in-virtue-of relation.

I contend that Hale's 'true in virtue of' relation that holds between propositions and natures must be primitive in the same way that Fine's TIVON relation is. One might think that the 'true in virtue of' relation is identical to some kind of consequence or truthmaker relation, but Hale's relation needs to be more subtle than that. For example, there are *prima facie* two ways that a proposition can be true in virtue of the nature of a thing (understood as the proposition bearing a relation to the nature, which is a property of the thing). Here are two examples:

1 – Socrates is a human.

2 – Socrates' nature is a property

1 is a good example of an essentialist claim about Socrates. 2 however is not an essentialist claim about Socrates, it is a claim about the nature of the nature. It is

²³¹ Fine (1995a), p.55, (1995b), p.275

²³² Fine (1995b), p.273

nothing to do with Socrates' essence, but everything to do with the metaphysics of properties. However, if Hale's relation is logical consequence or truthmaking then there can be no difference between 1 and 2. They are both true in virtue of the nature of Socrates (where 'the nature of Socrates' is taken to be a rigid designator for that property), and so are both essential to Socrates. This is not a satisfactory result.

One might tackle the problem by claiming that 2 is not true in virtue of the nature of Socrates, but in virtue of the nature of the nature of Socrates. As such, 2 would not be essential to Socrates, but to the nature of Socrates. This seems plausible, and avoids the present problem, but leads to an infinite regress, whereby each nature needs a nature of its own, and a nature of its nature, and a nature of the nature of its nature, *ad infinitum*. Whilst it does seem like the nature of Socrates should have a nature (after all, the nature is a property, and (according to Hale, at least) everything has a nature, including properties), to use this fact to try to explain 2 leads to regress. Overall it is safer to take the 'true in virtue of' involved in Hale's position as primitive. By claiming that 'true in virtue of' is primitive one can stipulate that 1 is true in virtue of the nature of Socrates, but not 2.

Third, as Hale suggests here and confirms elsewhere, he takes essentialist facts to be given by the definitions of *things*, not concepts.²³³ This is contrary to the Finean position, whereby the logical necessities are those propositions that bear TIVON to the logical concepts. Rather, Hale claims that logical necessities are true in virtue of the natures of logical entities. Further, Hale claims that the conceptual necessities can be accounted for in terms of the natures of the logical objects. More on this later. This has the advantage that it maintains a unified theory of necessity, and further, it seems perfectly compatible with Fine and Correia's work on essence.

²³³ Hale (2013), p.135, fn28, p.140, fn34

3.2 Hale on Necessity

Hale adopts a notation similar to Fine's. Read $\Box_{x_1 \dots x_n} P$ as the claim that P is true in virtue of the nature of n many objects. Hale presents the following explanation of necessity in terms of essence:

HALE: $\Box T$ because $\Box_{x_1 \dots x_n} S$, where S is true in virtue of the nature of $x_1 \dots x_n$, and T is a logical consequence of S .²³⁴

However, as Hale takes being true in virtue of the nature of to be closed under logical consequence, the better explanation

HALE': $\Box T$ because $\Box_{x_1 \dots x_n} T$

can be presented. Hale notes that the essentialist theory of necessity is “*structurally* parallel to those offered by [traditional] conventionalists, who likewise maintain that it is necessary that p when p 's truth is explained in a special way”.²³⁵

The account gives explanations of individual necessitation claims. These take the form $\Box T$ because $\Box_{x_1 \dots x_n} T$. For the second step one needs to give an account of metaphysical necessity in general. The individual explanation can be generalised:

HALE'': $\Box T$ iff $\exists x_1 \dots x_n (\Box_{x_1 \dots x_n} T)$

A proposition P is metaphysically necessary iff there are one or more objects such that P is true in virtue of the natures of those objects. For any well-defined class of entities that constitute a single kind there are propositions true solely in virtue of the natures of entities belonging to that class. For any such class of propositions there will be a corresponding kind of necessity. This has the consequence (as in Fine's theory) that all forms of necessity are special cases of metaphysical necessity. Arithmetical necessities are true in virtue of the natures of the numbers and arithmetic functions, and logical necessities are true in virtue of the natures of the logical functions. For instance, it is true that ‘necessarily, a conjunction of two

²³⁴ Hale (2013), p.146

²³⁵ Hale (2013), p.134 It is with no small amount of irony that I take inspiration from the neo-Aristotelian essentialist in providing the antirealist theory of necessity presented in Section Six. Convention imitates essence imitates convention.

propositions is true iff each conjunct is true' because it is true in virtue of the nature of conjunction that 'a conjunction of two propositions is true iff each conjunct is true'.

Hale takes conceptual necessities such as 'all bachelors are unmarried' to be necessary in the same way. Rather than identifying such claims as being true in virtue of the nature of concepts as Fine would have it, Hale insists that whilst what proposition is expressed by the sentence 'all bachelors are unmarried' may depend on the meanings of non-logical words or concepts, the proposition itself is true in virtue of the nature of logical entities, not concepts. Because the term 'bachelor' is defined as 'unmarried man' by what Quine would call a 'word sized convention', Hale takes the proposition expressed by 'all bachelors are unmarried' to be <anything that is both male and unmarried is unmarried>.²³⁶ The truth of this proposition is nothing to do with men or their marital status, but rather the logical entity that is conjunction.

Hale's theory of necessity does not rely on concepts, logical or otherwise. Whilst it does commit one to an ontology of logical entities like conjunction, at least such entities are better defined and understood, and the resulting account of necessity is more unified, not requiring the minimal necessity Fine introduces. Furthermore, there is no reason to think that the kinds of logical entities that Hale has in mind cannot be substituted into Correia's account in place of logical concepts. Indeed, the similarity leads me to believe that this may well be what Correia intended in the first place.

4. Complications with the Theory of Necessity

4.1 Sufficient Grounds

The central claim of the essentialist theory of necessity is that for any proposition P, $\text{nec}(P)$ iff some things exist such that P is true in virtue of the natures of those things. The question remains concerning certain propositions as to what they are true in

²³⁶ Quine (1936). This may be an oversimplification of the definition of 'bachelor', but not a problematic one.

virtue of the nature of. There are some propositions where there are no obvious candidates. For example the propositions <there are infinitely many natural numbers>, <if there is a thing then there is a singleton of that thing>, and <water has the elemental composition H_2O > are all necessarily true, and therefore must all be true in virtue of the natures of some things.²³⁷ Which things are they true in virtue of the nature of? It is not obvious that there are any finite pluralities of things that are up to the job, something more general is needed.

Hale suggests that we account for the necessity of such propositions by appealing not to the natures of objects, but the natures of properties. For example, <there are infinitely many natural numbers> is true not in virtue of the natures of all of the natural numbers, but in virtue of the nature of the property of being a natural number. A similar response can be given for <water has the elemental composition H_2O >, claiming that it is true in virtue of the nature of the property Water rather than any specific instances of water. However, in Chapter Four, Section Six I gave Correia's arguments for why such accounts would not work for the Finean. These arguments can be summarised as:

1. Ontological commitment to properties (whilst not a devastating criticism, Correia takes it to count against such a response).
2. Some predicates do not express properties, so not all generic essentialist claims can be explained by appeal to the objectual essences of properties (e.g. non-self-exemplifying properties, as such, are essentially properties).
3. Objects end up being ontologically dependent on the properties they essentially have (Socrates is not only essentially a man, but essentially has the property of being a man).²³⁸

In the place of explanations in terms of objectual essence, Correia insists that we adopt generic essence. However, by taking generic essentialist claims as primitive alongside objectual essentialist claims we sacrifice some of the simplicity of the theory, multiplying primitives in a way that is particularly problematic because those

²³⁷ These examples, and the related concerns they demonstrate come from Hale (2013), p.155, Cameron (2008), p.272, and Correia (2006), p.764, respectively.

²³⁸ Correia (2006), p.760-763

primitives are so very similar to each other but, being primitive, there is nothing to ground this similarity.

One may question how significant criticisms 1 to 3 are for Hale. Hale is already committed to properties, and as such is already committed to objects being dependent on those properties. If 1 and 3 are problematic for Hale's explanation of such propositions then it is only because they are already problematic for Hale's account in general. Furthermore, because of Hale's neo-Fregean understanding of properties, 2 is not relevant to the Hlean essentialist. According to Hale, for there to be a property P, all there needs to be is a corresponding meaningful predicate F. He claims that a property "just is a condition which things may or may not satisfy".²³⁹ As such, if there is a meaningful predicate, which Correia claims there is (and without which the argument collapses), then there is a property. As such, there is always a property corresponding to every predicate. This may be a satisfactory result for the strict Hlean essentialist. However, not everyone has a neo-Fregean conception of properties, and it is not clear that Hlean essentialism requires that one adopt one. What avenues of response are available to the Hlean essentialist who maintains a theory of properties that is susceptible to problem 2?

4.1.2 A First Response

As a first attempt at a response to problem 2, one might take an approach similar to that which Hale takes in accounting for conceptual necessity: arguing that such claims in fact express propositions that are true in virtue of the natures of logical objects. The sentence 'bachelors are unmarried' is true because (via word-length convention) 'bachelor' means unmarried man, and so the sentence expresses the proposition <anything that is both a man and unmarried is unmarried>. Perhaps the generic essentialist claim 'non-self-exemplifying properties are, as such, properties' expresses a more palatable proposition than it appears. The problematic claim could

²³⁹ Hale (2013), p. 165

be taken to express the proposition <anything that is a property and is not self-exemplifying is a property> $(\forall x((Fx \wedge \neg Rxx) \rightarrow Fx))$.

This reduces the generic essentialist claim ‘non-self-exemplifying properties are, as such, properties’ to the status of a triviality. If these were the only kinds of generic essentialist claims we could make about these kinds of properties then we might conclude from this that there are no significant essentialist facts here. However, we can make other claims that appear less susceptible to this kind of move. For example, ‘non-self-exemplifying properties are, as such, abstract’. This kind of response may be enough to account for the generic essentialist claim ‘non-self-exemplifying properties are, as such, properties’, but it cannot account for other claims attributing such properties with being abstract, being eternal, or being either instantiated or uninstantiated. Nor is it plausible that such a response can be generalised to all generic essentialist claims. For example, ‘water, as such, is H₂O’ cannot be accounted for in this way without making assumptions about the essence of water.

4.1.3 A Second Response

Hales’s use of the objectual essences of properties may appear to play right into Correia’s hands. However, there may be a response available based on how the mechanics of Hale’s account differ from those of Fine’s (upon which Correia’s original discussion is focused). One can distinguish between the property of being a natural number, and the Hale-property of the-nature-of-natural-numbers. Recall that unlike Fine’s TIVON relation, holding between a proposition and the thing itself, Hale posits a property of the thing, its nature, which the proposition is true in virtue of. In the case of the nature of a property, this is a second level property, non-identical to the property itself. Explaining generic essentialist claims in terms of the nature properties of properties is still an explanation in terms of the objectual essence of a property. Assuming that Correia is correct in claiming that there are predicates for which there are no corresponding properties (that is, assuming that one adopts Halean essentialism, but a different theory of properties), could there still be a nature

property for properties that are impossible in the way that the property of being a non-self-exemplifying property is? It seems that there is as good a case to say that there are such properties as there is to say that there is still a fact about what it is to be that thing.

There are two potential problems with this kind of response. First, such a nature property itself will be necessarily uninstantiated if there is necessarily no property that it is the nature of. Whether or not this is particularly problematic depends on one's metaphysics of properties, but it is not insignificant that it precludes one from accepting certain theories of properties, e.g. an Aristotelian theory of properties.

Second, since the property it is the nature of necessarily fails to exist, and the nature property is necessarily uninstantiated, even if one thinks that the nature property still exists it is no longer clear how it can be the nature of the property of being a non-self-exemplifying property. There is no link between the two. The property cannot instantiate anything because it cannot exist, so how exactly is the nature property supposed to be its nature in the same way that Socrates' nature is his nature (because he instantiates it), or even Socrates' merely possible twin sister (because she at least possibly instantiates it)? The only available option appears to be to take this relationship to be primitive.²⁴⁰ I do not find this response compelling. The primary motivation for an in-depth investigation into essence is so that it need not be taken as a mysterious primitive. Relying on this relationship (a relationship which, to my mind, should admit of further explanation) being primitive betrays that motivation.

4.2 Further complications with the Theories of Necessity

Here I present some further problems with the essentialist theories of necessity. How one might reply to these problems will play a significant role in determining the quality of the final realist theories as they will be presented in Section Five.

²⁴⁰ Plantinga (1974, 1976) does this.

The first problem for consideration is the apparent commitment to the existence of mere possibilities in Fine's theory. Fine's use of Π and Σ , quantifiers with the domain of all possibilities, in the object language of his logic might be a cause for concern. One might worry that if the theory is committed to the existence of mere possibilities, then this may mean that it relies on some kind of possible worlds ontology. If so then why even attempt to provide a theory of necessity in terms of essence when one has recourse to possible worlds? However, it seems that possibilities may be required for an essentialist theory of necessity. Not all necessities can obviously be accounted for without recourse to mere possibilities. For instance, we might ask how it is the case that Pegasus might have existed.²⁴¹

One might claim that the possibility of Pegasus' existence can be explained in terms of its compatibility with the essences of all existing things. A proposition P is necessary iff there is something that P is true in virtue of the nature of. Possibility can be defined in terms of necessity, so one can say that possibly P iff it is not necessary that not- P . Under the essentialist theory of necessity, a proposition P is possible iff it is not the case that there is something that not- P is true in virtue of the nature of. To use Hale's terminology, P is possible if there is no thing such that the nature of that thing *excludes* the truth of P . It is important to note that not excluding Pegasus in this context must mean something like 'not excluding something that matches our best description of Pegasus', as otherwise the account would be implicitly assuming that there is a Pegasus for things to exclude or not exclude. If the domain of the existential quantifier is only the actualia then this account is not fine-grained enough. Imagine there exists a contingent thing, the nature of which excludes the existence of Pegasus; call this thing anti-Pegasus. Suppose anti-Pegasus actually exists, and its existence (and nothing else's) is incompatible with Pegasus' existence, yet anti-Pegasus is contingent, so if anti-Pegasus did not exist then there would be nothing to stop Pegasus from existing. As such, a situation in which anti-Pegasus

²⁴¹ One might protest at this choice of example, claiming that either Pegasus does exist (as a fictional entity) or that it could not exist (for, given that Pegasus is fictional, there is nothing that it is to be Pegasus). If the reader finds this unconvincing, I ask that they substitute in the name of any specific merely possible object.

does not exist and Pegasus does exist is a possible one, contrary to the proposed explanation.

Alternatively, one might claim that the possibility of Pegasus' existence requires only that its existence be compatible with the essences of all the necessary things. This will not do either. Imagine that none of the necessary things exclude the existence of Pegasus, but that there is a class of contingent objects, each of which would exclude the existence of Pegasus, and that it is necessary that one object from that class must exist (though it is contingent which one). The world must contain one of the things that exclude Pegasus, even though none of those things exist necessarily. As such, it is not possible that Pegasus exist, contrary to the proposed account.

Alternatively again, one might claim that in order to be possible, Pegasus need only have a nature that does not exclude itself. That is, a nature that does not self-contradict. This too is inadequate. If anti-Pegasus were necessary, or if Pegasus were self-compatible, but incompatible with the laws of nature, then the existence of Pegasus should be impossible. However, simply by being non-self-contradictory, the proposed account renders Pegasus possible. This results in either a contradiction or the implausible result that despite being incompatible with the laws of nature, or a necessarily existent being, Pegasus is still possible. Further, it is not clear how this account could not presume the existence of (at least the merely possible) Pegasus.

There is further concern if we abandon mere possibilia. How can mere possibilia have essential profiles if there is no sense in which they exist? Under the Finean account, for something to have an essential profile it must be a possible relatum for the TIVON relation. Under the Halean account there must be a property that is the nature of that thing, yet as we saw in Section 4.1.3, it is not clear how a nature is related to a thing without it being instantiated (or at least possibly instantiated).²⁴²

At this point it may look like the conclusion to draw is that the inclusion of possibilia in the theory is a necessary, if slightly uncomfortable, part of the theory. If the inclusion of possibilia removes these problems, then perhaps it is justified.

²⁴² That is, unless we accept that the relationship between such properties and the things they are the natures of is primitive, à la Plantinga (1974, 1976).

Further, it does not seem that the use of possibilia requires any kind of possible worlds framework. However, it is not the case that it resolves the problems. If the domain of the existential quantifier in the explanation of necessity includes all possibilia, then it is not clear that the explanation is any more successful. If anti-Pegasus is possible, and the proposition <Pegasus exists> is possible iff it is not the case that there is something that <not <Pegasus exists>> is true in virtue of the nature of, and the domain of the quantifier includes all possibilia, then Pegasus' existence is not possible because it is excluded by (even the merely possible) anti-Pegasus. Even with the inclusion of possibilia, a more sophisticated theory is required in order to get the right results.

4.2.1 Hale to the Rescue?

In responding to some potential concerns that Hale presents, we can provide a response to the above problems. Hale highlights a further concern one might have about how anything can necessarily exist according to the essentialist theory of necessity. Consider the property P. As a property, P exists necessarily. For Hale, a property exists iff there could be a meaningful predicate corresponding to it. There could be such a predicate, and if we accept S5 then it is necessarily possible. As such, P exists necessarily.²⁴³ The obvious candidate to explain P's necessary existence is to claim that it is true in virtue of the nature of P that P exist. This seems an implausible explanation. As Hale says, it is "uncomfortably reminiscent of the notorious ontological argument for the existence of God."²⁴⁴ Hale suggests that instead of it being true in virtue of the nature of P that it exist simpliciter, it is true in virtue of the nature of P that it exist if certain conditions hold. That these conditions hold is a matter of necessity (and the source of this necessity is removed enough so as not to cause circularity), and thus P necessarily exists.

Hale's essentialist theory of modality also appears to be better equipped to respond to the other concerns. By developing his conception of essence in terms of

²⁴³ Hale (2013), p.176

²⁴⁴ Hale (2013), p.176 The argument he is referring to is that of Anselm (1078).

things having natures, Hale is in a position to evade commitment to possibilia in a way that he claims avoids the problems above. Hale expands upon his theory in discussing the following argument:

- 1 – It is true in virtue of the nature of Aristotle that he is a man.
- 2 – If it is true in virtue of the nature of Aristotle that he is a man then it is necessary that Aristotle is a man.
- 3 – Nothing can be true of Aristotle unless he exists.

Conclusion: Aristotle necessarily exists.

Hale responds to this by claiming that for there to be a nature there does not need to be a thing that it is the nature of.²⁴⁵ For instance, even if there were no water, there would still be such a thing as *what it is to be* water. By rejecting the assumption that there cannot be a nature of a thing unless that thing exists (a move that Hale is entitled to because the nature of a thing is non-identical to that thing, but that Fine is not entitled to make, as the TIVON relation must have the thing itself as a relatum) we can interpret the claim that Aristotle is essentially a man, not as entailing that Aristotle exists, but saying that being a man is (part of) what it is (or would be) to be Aristotle. That is, we can interpret the necessity that results from Aristotle essentially being a man, not as $\text{nec}(\text{Aristotle is a man})$, but as $\text{nec}(\forall x(x = \text{Aristotle} \rightarrow x \text{ is a man}))$. Interpreting it as such is sufficient to block the inference to Aristotle's necessary existence.

Fine presents a related problem that admits of the same solution.²⁴⁶ If it is necessary that Socrates is a man, but possible that he does not exist, then we can infer by modal axiom D (if necessarily P then possibly P) and conjunction introduction that it is possible that Socrates is a man and does not exist. However, it is not clear how Socrates could be a man and not exist. If we take Hale's advice and interpret the necessity that results from Aristotle essentially being a man as $\text{nec}(\forall x(x = \text{Aristotle} \rightarrow x \text{ is a man}))$ then this conclusion can be avoided. It is replaced with the much

²⁴⁵ Demonstrating further similarity between Hale's account and that of Plantinga (1974, 1976).

²⁴⁶ Fine (2005), p.328

more palatable claim that it is possible that Socrates does not exist and that nothing can be Socrates without being a man.

One might think that the natural direction to take this response in is to claim that natures exist necessarily, and as such we do not need to admit mere possibilia like Pegasus in order to ground modal facts about the Pegasus, as those facts are grounded in the nature of Pegasus. However, Hale does not claim that all natures exist necessarily. Some natures exist only contingently, and there might have existed some further natures in addition to those that already exist. Hale identifies which natures are contingent and shows that this does not have a detrimental effect on the theory.²⁴⁷

Purely general properties (properties corresponding to predicates with no embedded singular terms) exist necessarily. Contrast these with impure properties such as being a multiple of 4, being the brother of Aristotle, or being Alice or Bill. Impure properties depend on the things that feature in them for their existence. If that thing exists necessarily then so does the property; if that thing is contingent then so is the property. As such, natures that are purely general properties, and natures that are impure properties featuring something that exists necessarily, exist necessarily. Natures that are impure properties featuring something contingent are themselves contingent. Whilst the nature of Aristotle does not require the existence of Aristotle, there are contingent objects on which it does depend. Whilst some of his essential properties will be pure, such as being a man, some will be impure. For instance, if Aristotle's nature includes his biological origins then his nature features contingent objects, i.e. his parents. As such, Aristotle's nature is contingent because it depends for existence on Aristotle's contingent parents.

It is also through impure natures that the possibility of additional non-existent natures is realised.²⁴⁸ Since all possible purely general properties already exist, to say

²⁴⁷ Hale (2013), p.169

²⁴⁸ Hale (2013), p.166-167 gives an argument demonstrating that the purely general properties that exist are necessary, and necessarily all of the purely general properties that there are, and so there are no merely possible purely general properties. I do not include this argument here in full for the sake of brevity, but it hinges on Hale's claim that necessarily, a property P exists iff there could be a meaningful predicate of which P is the semantic value, combined with Hale's acceptance of S5 modal logic.

that something that does not exist might have existed is to say that there could have been one more thing of a certain already existing (though perhaps not actually instantiated) kind. Hale gives the example of a horse. For it to be possible that a ‘new’ horse exists is just for it to be possible that there be one more thing that is a horse and is not identical to any actual horse.²⁴⁹ Hale says:

If we were to assume that it is a possibility which can only be realized in this way, it is a purely general possibility – the possibility that there should exist an object or objects of some general kind. It is not, in other words, the possibility, concerning certain particular objects which do not to [*sic*] exist, that they might have existed – i.e. the possibility that certain merely possible objects should have been, not merely possible, but actual.²⁵⁰

As such, according to Hale’s view there are no merely possible objects, and so no *de re* modal claims about things that do not actually exist.

If necessity is grounded in the natures of things, and if there are some contingent natures, what effect does this contingency have on the theory of modality? Such a new horse would come with its own individual nature, and in turn there would be propositions that are true in virtue of (and excluded by) that nature, but Hale denies that this would place any additional constraint on what is metaphysically possible. Along with the new horse, there would also exist new propositions that would not otherwise have existed (e.g. propositions about that specific horse). The truth of some of these propositions would be excluded by the nature of the new horse (e.g. the proposition that the horse might have been an abstract object), but none of these would be propositions whose truth was not already excluded by what it is to be a horse in the first place.

²⁴⁹ For the sake of readability, and at the expense of accuracy, I follow Hale in using some temporal terms here. These terms should not be read literally, but metaphorically, as representing additions of ‘new’ objects that do not actually exist to the world, and subtractions of ‘old’ things that do actually exist.

²⁵⁰ Hale (2013), p.224

Conversely, if some natures that do exist had not existed, then some of the propositions that would have been true in virtue of those natures would not exist. Hale denies that this removes any constraints from metaphysical possibility. The risk is that there are some propositions whose truth are excluded by the natures of contingent things, but might be true if those things had not existed. This means that there would be fewer constraints on what is possible, thus allowing possibilities not accessible from the actual world. This is not the case because any proposition whose truth is excluded by the nature of an individual member of a kind also has its truth excluded by the general nature of that kind. Both of these results rely on the conditional form of essentialist claims that Hale advocates in his response to the earlier problem of things existing necessarily because of their essential profiles.

So what of Pegasus? The answer depends on whether Pegasus' nature exists. If so, the possibility of Pegasus would be realised by that nature being instantiated. If the nature does not exist (because it depends on some contingent thing that does not actually exist), then according to Hale, the possibility of Pegasus' existence is realised through the possibility of there being a thing that is of the kind Flying Horse (say) that matches the description given of Pegasus. I return to this in Section 4.2.3. However, there is a more problematic claim to consider. What of the claim that, had Aristotle not existed, it would have been possible that he exist?

If Aristotle had not existed (because his parents had not existed, and so neither had his nature), then Hale tells us that this would not affect what is or is not possible. However, there is a significant difference in that, had Aristotle not existed it would still have been possible that he exist. According to Hale there would be no *de re* possibility that Aristotle exist, as once we suppose that Aristotle's nature does not exist, the possibility of him existing *just is* the *de dicto* possibility of there being a new person, non-identical to all the existing people, who has enough of the features we associate with Aristotle.

It is important to note that under Hale's proposal it would not be the case that, had Aristotle not existed it would be impossible for him to exist, but that there would have been no modal facts about Aristotle at all. This is still a troubling conclusion.

The closest there would be to the possibility that Aristotle exist would be the possibility for there to be a person who is very similar, but not identical to, Aristotle. Call such a person Harry Stottle. Harry Stottle, whilst similar to Aristotle, is not enough to satisfy the *de re* claim that Aristotle could have existed.

Whilst this may be a consequence that Hale is willing to accept, there is a consequence which, come Hale or high water,²⁵¹ he will want to avoid. The falsity of the claim that, had Aristotle not existed, it would have been possible that he exist, appears to be a violation of the S5 axiom. Aristotle is a possible object, but he is not a necessarily possible object, because had his (contingent) nature not existed, Aristotle could not have existed.

However, there is a reply available to Hale. This objection relies not only on Aristotle not existing, but on Aristotle's nature not existing as well. For this to be the case I stipulated that Aristotle's parents, upon whom Aristotle's nature depends, not exist. However, this complicates matters. We should interpret the proposition under consideration as <Aristotle's parents did not exist and Aristotle exists>. The truth of this proposition is excluded by the general nature of what it is to be a human; humans just are things that cannot exist if their biological origins had not existed. As such, this proposition represents an impossibility, and rightly so. The possibility whose rejection jeopardised S5 was that Aristotle could have existed, even if he did not actually exist. This is true in the appropriate *de re* sense just so long as Aristotle's nature exists, and as Hale demonstrated earlier, the nature of Aristotle does not depend on Aristotle himself for its existence. In order for Aristotle's nature to not exist and allow the objection to go forward, Aristotle's parents must not exist, but to stipulate this is to negate the possibility being considered in the first place because the proposition under consideration would be the impossible <Aristotle's parents did not exist and Aristotle exists>.

²⁵¹ Sorry, couldn't resist.

4.2.2 Back to Pegasus

With all of this in place we are in a position to answer properly the question of the possibility of Pegasus' existence. We can modify the unsatisfactory essentialist explanation of necessity, according to which it is possible that Pegasus exist iff there are no things such that it is true in virtue of their nature that Pegasus does not exist. Consider the class of all natures, and consider every subclass of that class. A subclass is *allowed* iff it contains no nature that is excluded by a nature in the class (e.g. Pegasus and anti-Pegasus), it contains all of the natures of things that exist necessarily, and it contains natures to satisfy all necessary restrictions on how the world must be (e.g. if one of several contingent things must exist, then the class contains the nature of at least one of those things). For any *a*, it is possible that *a* exists iff there is an allowed subclass such that none of the natures in it exclude the existence of *a*. So for Pegasus:

PEGASUS: It is possible that Pegasus exist iff there is an allowed class of natures that do not preclude Pegasus' existence.

This can be generalised to:

POSSIBILITY: For any proposition *P*, possibly *P* iff there is an allowed subclass of the class of all natures such that none of the natures in it exclude the truth of *P*.

It is important to note that such a response is not available for the Finean, for whom there are no natures, or for someone who claims that all of the natures exist necessarily, for whom there will be only one allowed subclass (the improper subclass of all natures). It is also worth noting that this does not rescue de re modality concerning merely possible individuals for the Halean essentialist because there is no mere possibilium in question here, only a description of a de dicto possibility. The possibility of Pegasus is the de dicto possibility of something existing that satisfies our best description of Pegasus, not the possibility of a merely possible Pegasus' actuality.

An interesting consequence of such a position is that the explanations of necessity and possibility diverge from each other. Such an account as is given for

possibility would not give the correct results for necessity. Consider the corresponding definition of necessity:

NECESSITY: For any proposition P, necessarily P iff there is no allowed subclass of the class of all natures such that one of the natures in it excludes the truth of P.

Such a definition of necessity will not work under the Hlean account. Hale insists that the proper form of essentialist claims is conditional. Aristotle essentially being a man should in fact be interpreted as the claim that it is true in virtue of the nature of Aristotle that if anything is Aristotle, then that thing is a man. On its own this is not sufficient to secure the necessity of the proposition <Aristotle is a man>, there must also be something that is Aristotle. As such, necessity cannot be accounted for purely using the natures of things. In some cases at least, the things themselves must be involved as well.

4.3 Reductive Ambitions

Hale's account differs significantly from Fine's in terms of its reductive ambitions. Whilst Fine claims that his explanation of necessity in terms of essence constitutes a reduction of modality, Hale insists that such a claim misunderstands the kind of explanation that an essentialist theory is able to provide.²⁵² Hale's essentialist theory of necessity cannot be fully reductive because propositions about the natures of things are themselves necessary, and that necessity is, according to Hale, beyond explanation. The necessity of propositions that are true in virtue of the nature of a thing can be explained by appeal to the nature of those things (as per the essentialist theory of necessity). However, propositions *about* the nature of a thing, which are also necessary, cannot be explained by appeal to the nature of that thing without vicious circularity, nor can it be explained by appeal to the natures of other things (because that would undermine the claim that the proposition was true in virtue of the nature of the original thing). These are the only explanations available. Hale says:

²⁵² Hale (2013), p.158

[*T*]hat it is true in virtue of the nature of [*a* that *F*] admits of no further explanation. And if that is true (as I believe it is), it is necessarily so. But that necessity likewise admits of no further explanation.²⁵³

This stunts any reductive ambitions one might have because the source of the explanation of necessity is itself necessary, and that necessity is not subject to the same explanation. While Hale's account may be able to give explanations of individual necessities, it cannot give a reduction of modality itself to purely non-modal facts. However, Hale does not take the point of an essentialist theory of necessity to be to give a full reduction of modality, but rather to identify a base class of necessities that explain all of the others.²⁵⁴ This base class comprises those propositions that directly reflect the natures of things, and it is from this base class that all other claims necessitated inherit their necessity.

This conclusion is based on Hale's claim that absolute necessity cannot have its source in contingent facts. Hale's argument for this consists in arguing for the characteristic S4 principle, by which if $\text{nec}(P)$ then $\text{nec}(\text{nec}(P))$. However, Hale assumes that there is no way to honour S4 if the source of necessity is contingent. I hope that I have demonstrated in Chapter One that this is not the case. Having said this, showing that there is at least one way to honour S4 whilst claiming that the source of necessity is contingent is not the same as showing that this can always be done, and it is not clear how one might go about doing so in this case.

If we accept Hale's conclusion that his essentialist theory of necessity cannot be a reductive one, does this apply to the accounts of necessity as presented by Fine and Correia? Fine's theory of essence differs from Hale's in that instead of the 'true in virtue of the nature of' locution signifying a relation between a proposition and a property that is the nature of the thing, Fine takes it to signify that the primitive TIVON relation holds between a proposition and the thing itself. Presumably, to avoid getting the wrong result we have to take the necessities grounded in essentialist claims to be conditional in form, as Hale suggests. To say that Socrates is necessarily

²⁵³ Hale (2013), p.158

²⁵⁴ Hale (2013), p.158

a human simpliciter because he is essentially one is far too strong a result. If we take the relevant necessity to be that Socrates is a human if he exists, then the result is not too strong. However, if Socrates is contingent (which I assume he is) then he might not have existed. If the Finean essentialist wants to maintain their account of necessity then they had best find a way to make it such that Socrates' contingency does not jeopardise the account. This lends support to Fine's commitment to the existence of possibilia, including mere possibilia. This, combined with the conditional nature of the necessities that are grounded in essence, means that the propositions that are necessary because they bear TIVON to Socrates are necessary regardless of whether Socrates actually exists or not. As such, the Finean is committed to TIVON being a relation between propositions and not only actual things, but mere possibilia as well. If the Finean is committed to this then the modal status of those things is irrelevant to the account of necessity, and the Finean can maintain the reductive ambitions of their account.

It appears that the Correian essentialist, who takes generic essentialist claims to be primitive (either alongside or prior to objectual essentialist claims), must also sacrifice the reductive ambitions of their account. The generic essentialist claims are themselves necessary, and they too admit of no further explanation, and neither does their necessity. Just as Hale takes the point of the theory to be identifying a base class of necessities that provide a source of necessity for all other necessities, the generic essentialist claims provide such a base for the Correian essentialist (either alongside the objectual essences, or not, depending on whether one adopts Correia's reduction of the objectual to the generic).

5. The Choice of Essentialist Theories of Necessity

The realist essentialist is left with a choice of one of three positions. First, one can adopt a Finean position, where a is essentially F iff the proposition $\langle Fa \rangle$ bears TIVON to a . One can modify this position so that instead of relying on concepts, the logical and conceptual necessities are the result of the essences of logical entities. As

Correia shows, this position must adopt generic essence on a level at least as basic as objectual essence. This leaves the Finean essentialist with either two primitives, which is unappealing, or resorting to the reduction of objectual essence to generic essence. Taking the latter option means sacrificing the use of the TIVON relation, and making the potentially controversial move of having generic essences corresponding to haecceities. I claim that this move is problematic in Chapter Four, Section Six. The Finean is also committed to the existence of mere possibilia, which on top of the inclusion of logical objects makes for a particularly bloated ontology. However, this is the only position of the three that maintains its reductive ambitions for necessity. Though there are still problems with that reduction that are yet to be resolved.

Second, adopting Hale's essentialist theory of necessity commits one to abandoning the prospect of a reductive analysis of modality, instead providing an explanation of the necessity of most claims necessitated in terms of a (rather sizable) base class of essentialist facts. Having said this, it avoids the problems encountered by Fine in Section Four. Hale's essentialism also potentially manages to maintain the reduction of generic essence to the objectual essences of properties, thus maintaining some of the simplicity that the Finean was forced to surrender.

Third, the Correian essentialist is likewise forced to abandon any reductive ambitions, but in this regard is in no worse a position than the Hlean essentialist. The Correian does not struggle to account for generic essence, but is left with the same dichotomy as the Finean, to accept either having two essentialist primitives, or to resort to the potentially controversial reduction of objectual essence to generic essence.

As we have seen, whilst the realist essentialist has several viable theories of modality available to them, none are without their problems. In the next section I explore the potential for the antirealist essentialist to appropriate the realist essentialist strategy, and provide an essentialist theory of modality based on an antirealist theory of neo-Aristotelian essence.

6. An Antirealist Essentialist Theory of Necessity

6.1 Theory of Necessity

Following the realist essentialist theory of necessity, the antirealist can claim that a proposition P is necessarily true iff there is a thing x such that P is true in virtue of the nature of x . For the antirealist this translates to:

NECESSITY*: A proposition P is necessarily true iff there is a thing x such that P is the conclusion of a valid inference from a true GPI concerning x and relevant additional facts.

In what follows I address each of the challenges that arose for the realist theories presented earlier. I determine whether they apply to the antirealist theory, and in the case of those that do, examine the antirealist's options.

6.2 Concepts

Fine's realism accounted for logical and conceptual necessity using the essences of concepts. Little was said about what Fine meant by use of concepts, but Fine's introduction of weak necessity indicates that he did not intend it to mean some kind of platonic logical entities (as those used by Hale). This led to there being two different ways for propositions to be necessary, making the account less unified. In the interests of avoiding the diversion involved in ascertaining just what might be meant by concepts, I make the assumption that it is preferable for the antirealist to avoid talk of concepts if possible. Fortunately, it is possible. The antirealist can account for both conceptual and logical necessity without resorting to mysterious concepts.

Conceptual necessities can be accounted for in the same way that the antirealist was shown to account for generic essentialist claims in Chapter Four. A proposition is conceptually necessary iff it is the conclusion of a true GPI and appropriate

additional facts. Conceptual necessities of the form $\langle \text{all } Fs \text{ are } Gs \rangle$ ²⁵⁵ are true because Fs, as such, are essentially G. This is true because Fs are G, and because being G is a P-property (or the value of a P-category) for kind-K, which is a superkind of F. The relevant GPI is established by the following conventional practice:

CONVENTIONAL GENERIC GPI: If 'F' is a K-term then if G is the P-property of the things denoted by 'F', then 'F' applies to something in any possible situation only if it is G.

Each generic essentialist claim formed in this way is the source of a corresponding conceptual necessity, and for each conceptual necessity there is a generic essentialist claim that makes it necessary.²⁵⁶

Logical necessity can be accounted for in terms of the essences of logical objects. As in Chapter Two there are plausible GPIs that one can give for the various logical connectives. These rely on logical entities having the various properties that they do, and our GPIs then selecting those properly logical properties as being of importance.

GPI Conventional Schema: If ' \wedge ' is a logical connective-term then if ϕ is an inference licenced by the thing denoted by ' \wedge ', then ' \wedge ' applies to something in any possible situation only if it licences ϕ .²⁵⁷

That it is true in virtue of the nature of conjunction that a conjunction of two propositions is true iff each of the conjuncts is true means that that particular inference is one of conjunction's P-properties. Conjunction necessarily licences such an inference to be made on those grounds because we would not consider any instance where such an inference was not licensed to be an instance of conjunction. With this established we can account for logical necessity in the same way Correia does:

²⁵⁵ For example, 'all vixens are foxes', 'all bachelors are unmarried', and 'all drinks are potable'.

²⁵⁶ An equally viable alternative for the antirealist would be to take Hale's approach and reduce conceptual necessities to logical necessities and word-length conventions.

²⁵⁷ Note that this is different from the GPI presented in Chapter Two, as we are now equipped with the refinements from Correia.

LOGICAL: α is logically necessary iff for some plurality X of logical objects,
 $\beta(X) \vdash \alpha$.

A potential concern for this approach to logical necessity is that in avoiding the use of concepts one is committed to logical objects in the sense that Hale describes. One might think that this commits us to taking the no-TIVON robust-objects option for antirealist essentialism. After all, under the no-TIVON no-objects option there are no objects. This depends both on how we interpret logical objects, and how we interpret the no-objects part of the no-TIVON no-objects option. If we take no-objects to mean no robust objects whatsoever, and everything that is an object must have been the result of the conventional carving of the world, then the logical objects must be conventional in the same sense that humans and chairs are. It is not clear how such an account is one that contains logical objects rather than some kind of logical concepts. Alternatively, one might think that ‘no-objects’ only applies to physical objects, and allows for logical objects of the kinds Hale refers to. Logical objects are sufficiently different from the typical objects that are carved using the GPIs that this may be a plausible stance to take.

Whichever option the antirealist takes, they can still claim that the essential profile of the logical objects is due to the GPIs we hold concerning them. With the properly logical properties essential to them, the logical objects are sufficient to ground the logical necessities in the way they do in the realist theories. This can be done without reliance on concepts.

6.3 Derivative Essence

The antirealist account can accommodate derivative essence as suggested by Correia. The proposition $\langle\langle \text{Socrates is organic} \rangle \wedge \langle \text{the Eiffel Tower is synthetic} \rangle \rangle$ is irreducibly collectively essential to X (where X is a plurality containing Socrates, the Eiffel Tower, and conjunction), but further explanation can be given in essentialist terms. In the antirealist’s case this is because $\langle \text{Socrates is organic} \rangle$ is the conclusion of a valid inference from one of Socrates’ GPIs and the empirical fact that he is

organic; <the Eiffel Tower is synthetic> is the conclusion of a valid inference from one of the Eiffel Tower's GPIs and the empirical fact that it is synthetic; and it is the conclusion of a valid inference from a GPI for conjunction and a relevant additional fact that when two propositions are true, so is the conjunction of those two propositions.

The antirealist can explain that <<Socrates is organic>^<the Eiffel Tower is synthetic>> is derivatively essential to X in the same way that Correia says the Finean can:

DERIVATIVE: α is derivatively essential to X iff α does not belong to $\beta(X)$ and $\beta(X) \vdash_{\log(X)} \alpha$.

Allowing the antirealist the inferential nature the essences of the logical objects means that this move is open to them, just as it is the realist.

6.4 Commitment to Mere Possibilia

By showing that, through the conventionalist mechanism, things can have essences without existing, the antirealist can provide an account of possibility that does not rely on ontological commitment to mere possibilia.

Just as in Hale's account, where natures can exist even if the thing that they are a nature of does not, we can hold GPIs about things that do not exist. This means that for something to have an essence does not require that it exists. For instance, we can hold GPIs concerning Pegasus. The additional facts that we need for these GPIs to generate essentialist claims come from the myth of Pegasus. They are not empirical facts, but as we established in Chapter Two, they need only be from a source external to the GPI.

Like Hale's properties, for there to be a GPI requires only that there be a meaningful predicate. This follows from the relationship that Fine posits between essence and meaning. According to Fine, definition of the meanings of terms is a special case of real definition for things. As such, if there is a meaningful predicate then there must at least be an essence that is the meaning of that term. If we accept

S5, and there is a possible predicate, then there is necessarily a possible predicate. Since this is the condition required for there to be GPIs for a kind, the kinds that exist do so necessarily. Further, since necessity entails actuality, any that could exist (and thus must exist), do exist.²⁵⁸ As such, for there to be a set of GPIs governing something's essence does not require that that thing exist, only that there is a meaningful predicate of what it is to be that thing.

However, due to the conventional nature of the GPIs, it must also be the case that we could just as easily have formed other kinds from those that we had, or that we could have failed to have formed some of those we did. This apparent tension is addressed by my discussion of the contingency problem in Chapter One. As we saw there, by distinguishing between the conventional practices we hold, and the carvings they ground, we can avoid the problem. The necessity grounded in the GPIs for kinds is sufficient to demonstrate that they are all of the possible kinds *relative to the carving*.

Having established that we can hold GPIs that ground the essences of things that do not exist, we can address the issue of accounting for their possibility. First, we can identify three ways that something can be impossible:

1. Some necessarily existing plurality *X* excludes *P*.
2. It must be the case that there exists some plurality *X* (where at least some of the members of *X* are contingent) that excludes *P*.
3. *P* somehow excludes itself (for example if *P* is the statement that *a* exists, and it is true in virtue of the nature of *a* that *a* does not exist).

Any account of the possibility of some proposition *P* will have to exclude 1-3. We can do this by saying that possibly *P* iff it is not the case that there must exist a plurality *X* such that not *P* is true in virtue of *X*'s nature. Because of how the essentialist theory of necessity works, there must be an *X* iff there is a *Y* (which may or may not be identical to *X*) such that it is true in virtue of the nature of *Y* that *X* exists. This is expanded to:

²⁵⁸ Hale (2013), p. 166-167 gives such an argument for the purely general properties.

Possibly P iff it is not the case that there exists a plurality *X* such that it is true in virtue of the nature of *X* that there exists a plurality *Y* such that it is true in virtue of the nature of *Y* that not P.

In full antirealist presentation, this looks like:

ANTIREALIST POSSIBILITY: Possibly P iff there does not exist an *X* such that it is the conclusion of a valid inference from *X*'s GPIs and relevant additional facts that there exists a *Y* such that that it is the conclusion of a valid inference from *Y*'s GPIs and relevant additional facts that not P.

Applying this to the case of Pegasus' possibility we can say that the mere possibility of anti-Pegasus does not affect the possibility of Pegasus because it is not the case that anti-Pegasus (or something else that would exclude Pegasus) must exist. If it is the case that something that excludes Pegasus must exist then Pegasus is impossible, and rightly so.

Whilst this account does not explicitly rule out P self-excluding, this can be interpreted as P being excluded by either the nature of the logical objects, or the generic essentialist claims. For instance, propositions asserting that something instantiates the property of being a non-self-exemplifying property would be excluded by the natures of the logical objects, whereas the proposition that something is both green and red all over would be excluded by the generic essentialist facts about colour.

Is this available to the Finean? No. For the Finean the essence of an object is dependent on that object existing (even if only as a mere possibilium) so that it can be a relatum for the TIVON relation. As such, if things like anti-Pegasus do not exist, their possible existence cannot be brought to bear because without them existing in some way there is nothing that it is to be anti-Pegasus.²⁵⁹ Imagine that there had to be something that excluded Pegasus, but that every possible thing that excludes Pegasus is itself contingent. This is a situation in which Pegasus' existence would be

²⁵⁹ Again, whilst there may be a precedent to deny this from Plantinga, such an option is *i*) (I claim) contrary to the true motivations behind a theory of essence, and *ii*) not available to the Finean essentialist.

impossible, but for the Finean there would only be essentialist facts about the anti-Pegasus that actually exists, and since that is contingent, Pegasus would be possible, which is the wrong result. This is not a problem for the antirealist because an object does not need to exist for us to hold GPIs about it.

A consequence of this account is that, as in Hale's account, there can be no *de re* modality for things that do not exist. Whilst there are GPIs governing the kinds that things are members of, and even though we could stipulate what additional facts might supplement those GPIs, that would not be enough to single out a specific *de re* possibility. This is because essence cannot, contrary to Hale, do the job so distinguishing a thing from every other possible thing. We could specify all of the appropriate supplementary facts for the GPIs concerning Socrates to give him a possible sister, but the best we could do is specify that there is a new thing that would satisfy the conditions of being Socrates' sister. Such a sister would not be a specific thing that we are saying exists possibly, as one would with a genuine *de re* possibility.

6.5 Reductive Ambitions

The antirealist position does not run the same risk to its reductive ambitions as the Hlean and Correian theories of necessity do. The grounds of necessity are in contingent conventions about kinds and individuals. Whilst the essentialist facts grounded by the conventions are necessarily true, unlike in Hale's theory this necessity *does* admit of further explanation. Those essentialist facts are necessary because they are the conclusion of a valid inference from GPIs and relevant additional facts. Whilst those additional facts may well be necessarily true, their necessity also admits of further explanation. Ultimately all of the necessity is grounded in the GPIs. These are contingent, and so require no further explanation. Concerns relating to the contingency of the GPIs can be addressed in the same way presented in Chapter One.

7. Conclusion

In this chapter I presented the Finean, Correian, and Halean options for the realist essentialist theory of necessity. I explored various challenges that those positions faced, and showed to what degree they can be overcome. I concluded that each is left with problems. I do not claim that these problems cannot be overcome, only that they must be responded to if these positions are to remain competitive.

I presented an antirealist essentialist theory of necessity, developed from the antirealist theory of essence as presented in Chapter Four, which appropriates the realist strategy to some extent. I then demonstrated that such a position is either immune to, or in a better position to respond to, each of the challenges that face its realist counterparts.

Conclusion

To conclude this project, I show how each of its three goals have been achieved, and suggest further work to be done in the future.

1. Goals

This project had three goals. The first goal was to show that in the debate between modal essentialism and neo-Aristotelian essentialism the Finean arguments are not decisive against the modal essentialist, but are suggestive of a different reason to favour neo-Aristotelian essentialism. This was achieved in Chapter Three in two ways. First, by taking the arguments that Fine gave against the modal essentialist, and considering them in terms of the proposed interpretation of the debate I suggested in the introduction to this thesis. Second, by presenting a form of modal essentialism that promises to provide an account of essence on Fines terms whilst avoiding Fine's arguments.

My proposed interpretation of the debate between modal essentialism and neo-Aristotelian essentialism claimed that the rival accounts do not posit different explanations for something well-understood and uncontroversial, but rather they propose different notions of essence that the incomplete data of our pre-philosophical essentialist opinions would be compatible with. How we choose a theory of essence should then be a balance of theoretical virtues and wider metaphysical considerations. By insisting that the modal essentialist be judged against the standard of extensional accuracy to pre-philosophical opinions, Fine makes two mistakes. First, he assumes that there is such an extensive base of pre-philosophical opinions that potential theories must adhere to. Second, he assumes that this base coincides with the neo-Aristotelian conception of essence. In light of this discussion it is clear that at best the Finean arguments fail to decisively defeat modal essentialism, and at worst they beg the question against it. However, if we adopt my proposed understanding of the debate, it is plausible that neo-Aristotelian essentialism still

comes out on top. This is because it fits better into our wider metaphysical theorising, providing insight into the link between modality and ontology through a theory of necessity and ontological dependence, and, on the antirealist neo-Aristotelian account, providing a better explanation of why we essentialise in the first place.

The suggested version of modal essentialism in Chapter Three is a modification of inspired by the way that Fine develops his own neo-Aristotelian position. One might call such a position *truthmaker modal essentialism*. This claimed that a proposition P is part of *a*'s constitutive essence iff P is true at every world at which *a* exists, and at every world where *a* exists, *a* and *a* alone is the thing that makes P true. Using this definition for the constitutive modal essence, the truthmaker modal essentialist can respond to Fine's arguments. Even assuming the Finean interpretation of the debate between modal and neo-Aristotelian essentialism, truthmaker modal essentialism is able to avoid Fine's arguments. As such it is a real competitor with Fine's position, even by his own lights. The presentation of the theory was brief, and there are challenges that a full development of the position would need to overcome. For example, for any object that makes a proposition P true in every world in which it exists, it seems that for some propositions (though not all) it may also be the case that the proposition is made true by that object's parts as well, and as the theory stands, if two things make the proposition true then it is essential to neither. It is also not clear how the position would be developed to properly incorporate generic essentialist statements. However, there is potential in the position and it warrants further work.

The second goal was to demonstrate that an antirealist theory of essence is a viable position to hold, and one worthy of wider consideration in contemporary metaphysics. This was demonstrated throughout the project by showing the various options available to the would-be antirealist essentialist and further, demonstrating that there are some distinct advantages that these positions have over their realist counterparts (beyond the predictable virtues of parsimony). For example, in Chapter Four I demonstrated that the antirealist neo-Aristotelian essentialist can better account for generic essentialist claims. Whereas the Finean essentialist has to adopt a

new primitive, or engage in a reduction of objectual essence to generic essence (a reduction that is potentially controversial), the antirealist is able to account for generic essence using only the tools already available to them. Furthermore, in Chapter Five I demonstrated that the antirealist neo-Aristotelian essentialist can provide an essentialist theory of modality that is seemingly no worse than that produced by its realist counterparts, and further, that is in a better position to reply to the problems that beset its competitors.

The third goal was to show that an antirealist theory of essence is a viable position regardless of whether one interprets essence as the modal essentialist or the neo-Aristotelian essentialist does. I demonstrated that this is the case by developing an antirealist account for each of the two different interpretations of essence in Chapters Two and Four. Whilst it became apparent that the neo-Aristotelian antirealist essentialism is the more promising theory, this is purely in virtue of the neo-Aristotelian interpretation of essence being more promising in the first place. Regardless, both the modal and neo-Aristotelian interpretations were shown to have viable antirealist accounts available to them based on the Sidelleian conventionalist machinery.

2. Future work

Having completed the task that I set out to do in the introduction to this project, there are various further tasks that present themselves as worthwhile continuations of the work completed here. The nature of this project, as a demonstration that antirealism about essence is a viable option worthy of discussion, by its very nature means that there is much more to be said. This project did not aim to argue that antirealist essentialism is the best theory of essence, or that there is anything wrong with realist theories of essence (though some points of comparison between the realist and antirealist positions did arise during the discussion). As such, there is more work to be done in further developing antirealist essentialism as a theory, critiquing and defending it, and investigating its applications in wider philosophical

theorising, than I can list here. Below I list three tasks which are warranted based on the work done in this project. However, this should not be taken as the limit on future work to come from this thesis.

The dismissal of the deflationary and Humean neo-conventionalist positions was motivated by their lack of explanation as to what the relevant conventions involved in grounding necessity were for, and how those conventions should be able to achieve their ends. The accounts did not present any kind of practical motivation for the adoption of the relevant conventions, and this was one of the requisites for such a theory as presented in the introduction to this thesis. However, the lack of such detail does nothing to say that this kind of additional explanation could not be given. There is nothing to stop these positions from being developed along different lines than Sidelleian neo-conventionalism. If such additional work were to be completed, there may well be alternative theories that would make viable candidates for antirealism about essence. Such a development would be a worthwhile project, as it would establish whether the Sidelleian mechanism is the only game in town, or if it needs to compete with antirealist alternatives as well as realist ones.

In addition, at the end of Chapter Three I briefly presented the germ of a theory I have labelled truthmaker modal essentialism. I contend that such a theory of essence has promise, and is worthy of further attention, which would have been a diversion in the context of the current project. There are certain challenges that such a position must contend with. For instance, it is not immediately clear how such an account would explain essential relations (as they involve more than one object), or essentialist claims relating to properties, or generic essentialist claims. A further investigation into the advantages and disadvantages of such a position might render a new modal essentialist theory that is competitive with its neo-Aristotelian alternatives. Furthermore, whilst the position as hinted at in Chapter Three was presented in a realist form, there is no immediate barrier to an antirealist interpretation of truthmaker modal essentialism. This could potentially present a new form of antirealist essentialism, and would be worthy of investigation.

Most notably, Chapter Five gave only a preliminary investigation of the potential for essentialist theories of necessity (indeed, a full investigation of such theories would require a thesis in itself). The need to present and evaluate multiple realist essentialist theories of necessity, so that the antirealist account might have a worthy basis for comparison, meant that the discussion lacked the depth that would be needed to fully evaluate the potential of essence to provide an explanation of necessity. The appropriate next step to continue this research would be to conduct a more rigorous examination of the full array of options for such positions, so that a proper survey of essentialist theories of necessity might be provided. Only when this is done satisfactorily will it be possible to go into similar depth concerning the prospects on an antirealist essentialist account of necessity. As it stands, the results of Chapter Five only go so far as to suggest that such an account is feasible and worthy of further consideration, and that there are some *prima facie* reasons to think that such an account may be competitive with its realist counterparts.

3. Conclusion

Having satisfied the goals of this project, and highlighted future work to build on its results, I conclude that antirealist essentialism is indeed a viable theory of essence that is very much worthy of consideration in contemporary debate, and that this result is correct regardless of whether one interprets essence as the neo-Aristotelian or the modal essentialist does (and indeed, that consequently the debate between modal and neo-Aristotelian essentialism is not settled for the realist either).

Bibliography

- Anselm, Saint, Archbishop of Canterbury. 1078/1974: *Proslogion*. In Hopkins, J. & Richardson, H. (eds. & trans.) *Anselm of Canterbury* Vol 1. London: SCM Press Ltd.
- Ayer, A. J. 1936. *Language, Truth and Logic*. London: Gollancz, Ltd.
- Ayer, A. J. Whiteley, C. H. & Black, M. 1936b. Truth by Convention: A Symposium. *Analysis* 4(2-3), pp.17-32.
- Blackburn, S. 1984. *Spreading the Word*, Oxford: Clarendon Press.
- Blackburn, S. 1993. Morals and Modals. In Blackburn, S. *Essays in Quasi-Realism*. Oxford: Oxford University Press, pp.52-74.
- Cameron, R. 2008. Truthmakers and Modality. *Synthese*. 164(2), pp.261-280.
- Cameron, R. 2009. What's Metaphysical about Metaphysical Necessity? *Philosophy and Phenomenological Research* 79(1), pp.1-16.
- Cameron, R. 2010a. On the Source of Necessity. In Hale, B. & Hoffman, A. (eds.), *Modality: Metaphysics, Logic and Epistemology*. Oxford: Oxford University Press.
- Cameron, R. 2010b. The Grounds of Necessity. *Philosophy Compass* 5(4), pp.348-358.
- Carnap, R. 1937. *The Logical Syntax of Language*. London: Paul, Trench, Trubner & Co.
- Carnap, R. 1947. *Meaning and Necessity*. University of Chicago Press.
- Carroll, L. 1895. What the Tortoise said to Achilles. *Mind* 4(14), pp.278-280.
- Chalmers, D.J. 2002. Does Conceivability Entail Possibility? In Gendler, T.G. & Hawthorne, J. (eds.), *Conceivability and Possibility*. Oxford University Press, pp.145-200.
- Correia, F. 2006. Generic Essence, Objectual Essence, and Modality. *Noûs* 40(4), pp.753-767.

- Correia, F. 2007. (Finean) Essence and (Priorean) Modality. *Dialectica* 61(1), pp.63–84.
- Correia, F. 2012. On the Reduction of Necessity to Essence. *Philosophy and Phenomenological Research* 84(3), pp.639-653.
- Della Rocca, M. 1996. Essentialism: Part 2. *Philosophical Books* 37(2), pp.81-89.
- Divers, J. & Elstein, D. 2012. Manifesting Belief in Absolute Necessity. *Philosophical Studies* 158(1), pp.109-130.
- Divers, J. & González-Varela, J.E. 2013. Belief in Absolute Necessity. *Philosophy and Phenomenological Research* 87(2), pp.358-391.
- Dummett, M. 1959. Wittgenstein's Philosophy of Mathematics. *Philosophical Review* 68(3), pp.324-348.
- Einheuser, I. 2006. Counterconventional Conditionals. *Philosophical Studies* 127(3), pp.459-482.
- Elder, C.L. 2006. Conventionalism and Realism-Imitating Counterfactuals. *Philosophical Quarterly* 56(222), pp.1-15.
- Fine, K. 1994. Essence and modality. *Philosophical Perspectives* 8, pp.1-16.
- Fine, K. 1995a. Senses of Essence. In Sinnott-Armstrong, W. Raffman, D. & Asher, N. (eds.), *Modality, Morality and Belief. Essays in Honor of Ruth Barcan Marcus*. Cambridge: Cambridge University Press.
- Fine, K. 1995b. Ontological Dependence. *Proceedings of the Aristotelian Society* 95, pp.269-290.
- Fine, K. 1995c. The Logic of Essence. *Journal of Philosophical Logic* 24(3), pp.241-273.
- Fine, K. 2000. Semantics for the Logic of Essence. *Journal of Philosophical Logic* 29(6), pp.543-584.
- Fine, K. 2005. *Modality and Tense*. Oxford: Oxford University Press.
- Fine, K. 2007. Response to Fabrice Correia. *Dialectica* 61(1), pp.85–88.

- Frege, G. 1884/1980. *Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number*, Translated by Austin, J. L. Northwestern University Press.
- Goodman, N. 1983. *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.
- Gorman, M. 2005. The Essential and the Accidental. *Ratio* 18(3), pp.276–289.
- Hale, B. 2002. The Source of Necessity. *Noûs* 36(s16), pp.299-319.
- Hale, B. 2013. *Necessary Beings: An Essay on Ontology, Modality, and the Relations Between Them*. Oxford: Oxford University Press.
- Kant, I. 1781/2007. The Impossibility of an Ontological Proof of the Existence of God. In Kant, I. *Critique of Pure Reason*. Basingstoke: Palgrave Macmillan, pp.500-507.
- Kaplan, D. 1989. Demonstratives. In Almog, J. Perry, J. & Wettstein, H.K. (eds.), *Themes From Kaplan*. New York: Oxford University Press, pp.481-565.
- Kripke, S.A. 1971. Identity and Necessity. In Munitz, M.K. (ed.), *Identity and Individuation*. New York University Press, pp.135-164.
- Kripke, S.A. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Lewis, D. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Lewis, D. 1983. New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61, pp.343-377.
- Lewis, D. 1986. *On the Plurality of Worlds*. Oxford: Blackwell Publishers.
- Lewy, C. 1976. *Meaning and Modality*. Cambridge: Cambridge University Press.
- Locke, J. 1689. *An Essay Concerning Human Understanding*. Menston: Scolar Press.
- Lowe, E. J. 2008. Two notions of being: Entity and essence. *Royal Institute of Philosophy Supplements* 83(62), pp.23-48.
- Mackie, P. 2006. *How Things Might Have Been: Individuals, Kinds, and Essential Properties*. New York: Oxford University Press.

- Oderberg D.S. 2007. *Real Essentialism*. London: Routledge.
- Plantinga, A. 1974. *Nature of Necessity*. Oxford: Clarendon Press.
- Plantinga, A. 1976 Actualism and Possible Worlds. *Theoria* 42(1-3), pp.139-160.
- Putnam, H. 1973. Meaning and reference. *Journal of Philosophy* 70(19), pp.699-711.
- Putnam, H. 1975. The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science* 7, Minneapolis: Minnesota University Press, pp.131-193.
- Quine, W. V. O. 1936. Truth by Convention. In Gibson, R.F. (ed.) *Quintessence: Basic Readings from the Philosophy of W.V. Quine*. Cambridge, MA: Harvard University Press, pp.3-22.
- Quine, W. V. O. 1951. Two Dogmas of Empiricism. *Philosophical Review* 60(1), pp.20-43.
- Quine, W. V. O. 1953. Three Grades of Modal Involvement. In Gibson, R.F. (ed.) *Quintessence: Basic Readings from the Philosophy of W.V. Quine*. Cambridge, MA: Harvard University Press, pp.379-399.
- Quine, W. V. O. 1960. Carnap and Logical Truth. *Synthese* 12(4), pp.350-374.
- Russell, G. 2008. *Truth in Virtue of Meaning*. Oxford: Oxford University Press.
- Russell, G. 2010. A New Problem for the Linguistic Doctrine of Necessary Truth. In Wright, C.D. & Pedersen, N.J.L.L (eds.), *New Waves in Truth*. Palgrave Macmillan, pp.267-281.
- Shalkowski, S.A. 2008. Essence and Being. *Royal Institute of Philosophy Supplements* 83(62), pp.49-63.
- Shalkowski, S.A. Forthcoming. Essence and Nominalism. In Dumitru, M. (ed.), *Metaphysics, Meaning, and Modality. Themes from Kit Fine*. Oxford: Oxford University Press.
- Sidelle, A. 1989. *Necessity, Essence, and Individuation: A Defense of Conventionalism*. Cornell University Press.

- Sidelle, A. (1998). A Sweater Unravelled: Following One Thread of Thought for Avoiding Coincident Entities. *Noûs* 32(4), pp.423-448.
- Sidelle, A. 2009. Conventionalism and the Contingency of Conventions. *Noûs* 43(2), pp.224-241.
- Sidelle, A. 2010. Modality and Objects. *Philosophical Quarterly* 60(238), pp.109-125.
- Sider, T. 2003. Reductive Theories of Modality. In Loux, M.J. & Zimmerman, D.W. (eds.), *The Oxford Handbook of Metaphysics*. Oxford University Press, pp.180-208.
- Sider, T. 2011. *Writing the Book of the World*. Oxford: Oxford University Press.
- Van Cleve, J. 1994. Descartes and the Destruction of the Eternal Truths. *Ratio* 7(1), pp.58-62.
- Wiggins, D. 1976. The De Re 'Must': A Note on the Logical Form of Essentialist Claims. In Evans, G. & McDowell, J.H. (eds.) *Truth Meaning and Essays in Semantics*. Oxford: Clarendon Press, pp.285-312.
- Wildman, N. 2013. Modality, Sparsity, and Essence. *Philosophical Quarterly* 63(253), pp.760-782.
- Wildman, N. Forthcoming. How (Not) to be a Modalist about Essence. In Jago, M. & Barker, S. (eds.) *Reality Making*.
- Wisdom, J. 1938. Metaphysics and verification (I.). *Mind* 47(188), pp.452-498.
- Wright, C. 1980. *Wittgenstein on the Foundations of Mathematics*. Cambridge, MA: Harvard University Press.
- Wright, C. 1985. In Defence of the Conventional Wisdom. In Hacking, I. & Lewy, C. (eds.) *Exercises in Analysis: Essays by Students of Casimir Lewy*. Cambridge: Cambridge University Press.
- Wright, C. 1992. *Truth and Objectivity*. Cambridge, MA: Harvard University Press.
- Zalta, E.N. 2006. Essence and modality. *Mind* 115(459), pp.659-693.

